Technische Universität München
Fakultät für Elektrotechnik und Informationstechnik

Microscopic, Mesoscopic, and Macroscopic Learning Algorithms and Effects in Multi-Agent Systems

Ezra Tampubolon

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr.-Ing. Wolfgang Kellerer

Prüfer der Dissertation:

1. Prof. Dr.-Ing. Dr.rer.nat. Holger Boche
2. Prof. Dr.-Ing. Eduard A. Jorswieck

Die Dissertation wurde am 10.09.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 25.11.2021 angenommen.

Abstract

Complex large-scale networked systems are incorporated in widespread structures in our society. Prominent examples are the economic market, traffic systems, and the internet. Moreover, they constitute the foundations of several important current and future applications. Examples are the Internet-of-Things, wireless sensor networks, supply-chain management, cooperative robotics, and many more. The challenge involved with the emergence of such systems is to understand their dynamics and to coordinate their components toward achieving a desired global objective. In doing so, we consider three different social levels: First, the macroscopic level referring to the system as a whole; Second, the mesoscopic level relating to the clusters of the components; Third, the microscopic level relating to the individual components.

In the first part of this thesis, we focus on the setting of selfish learning agents from the game-theoretic point of view. Assuming that the underlying game is aggregative, we study the extent to which the lack of agents' knowledge about the global change of the populations' state is acceptable for the population of learning agents so that it still reaches the underlying equilibrium state - Wardrop equilibrium. Subsequently, we design a decentralized pricing mechanism providing the agents incentives for resource sustainable behavior. Using tools from the variational analysis and martingale theory, we provide mechanism parameter choices leading the population to the generalized variational equilibrium of the underlying game with the coupled constraints. In doing so, our emphasis is on the non-asymptotic guarantees.

In the second part of the thesis, we restrict our view to a single learning agent. Driven by the aim to solve the problem of online learning with (resource) constraints, we introduce a new aggregate constraint violation measure called h-CFit. This measure motivates us to design an efficient online learning method minimizing the regret and having a tighter constraint violation guarantee then state of the art.

In the third part of the thesis, we consider the problem of locally-cooperative control of a multi-agent system. At first, we concern with distributed optimization in a networked system. Utilizing the theory of stochastic differential equations, we analyze the continuous-time distributed subgradient algorithm by deriving a non-asymptotic high probability and expectation bound for the distance of the average function value of the iterates of the agents providing suggestions of parameter choices leading to the algorithm's success. The rest of the last part of the thesis is devoted to the investigation of the cluster behavior of the opinion of a population of informational biased communicating

agents. We introduce the novel notion of substochastic complementation, which gives an efficient way to part the network into clusters and allows us to quantify the degree of the clusterness of the agents' opinion.

Zusammenfassung

Komplexe, im großen Maßstab vernetzte Systeme sind Bestandteil vieler Strukturen in unserer Gesellschaft. Prominente Beispiele von solchen Strukturen sind die Marktwirtschaft, das Internet und die Verkehrssysteme. Ebenfalls, bilden solche Systeme die Grundlage vieler gegenwärtiger und zukünftiger Anwendungen, die einen großen Einfluss auf unser Leben haben, bzw. haben werden. Beispiele sind Internet der Dinge, drahtlose Sensornetzwerke, Versorgungskettenmanagement, kooperative Roboter und vieles mehr. In dieser Arbeit, beschäftigen wir uns mit zwei Herausforderungen, die mit solchen Systeme auftreten: das Verstehen des zugrunde liegenden Dynamik und die Koordinierung der Komponenten zu einem gewünschten globalen Zielzustand. Dabei betrachten wir drei hierarchische Systemebenen: Die makroskopische Ebene, die das System als Ganzes beschreibt; die mesoskopische Ebene, die die einzelnen Gruppen der Komponenten beschreibt; die mikroskopische Ebene, die die individuelle Komponente in Detail betrachtet.

In dem ersten Teil dieser Arbeit konzentrieren wir uns auf die Konfiguration der lernenden Agenten aus spieltheoretischer Sicht. Mit der Annahme, dass das zugrundeliegende Spiel aggregativ ist, untersuchen wir, in welchem Maße Agentenunwissenheit über globale Veränderungen des Populationszustands erlaubt ist, sodass das System der Agenten das Equilibrium des Spiels (Wardrop Equilibrium) dennoch erreicht. Nachfolgend entwerfen wir für allgemeine Spiele einen dezentralisierten Preismechanismus, der den Agenten Anreize für Ressourcen-nachhaltiges Verhalten gibt. Wir liefern Parameterwahlen, die die Population zu dem generalisierten Nash Equilibrium des zugrundeliegenden Spiels mit gekoppelten Beschränkungen führen. Dabei nutzen wir Werkzeuge der Variationsanalysis und der Martingaltheorie. Unser Schwerpunkt liegt auf nicht-asymptotischen Garantien.

In dem zweiten Teil der Arbeit beschränken wir unsere Sicht auf einen einzelnen Agenten und betrachten dabei das Problem dessen Lernens mit Nebenbedingungen. Dabei führen wir ein neues Maß für seine aggregierten Verletzungen, genannt h-CFit, ein. Dieses Maß erlaubt uns eine effiziente Methode für das Onlinelernen bereitzustellen. Die Methode minimiert die Reue des Lernenden und bietet eine Garantie für Bedingungsverletzungen, die präziser im Vergleich zu anderen aus der aktuellen Forschung ist.

Der dritte Teil dieser Arbeit ist dem Problem der lokal-kooperativen Steuerung eines Mehragentensystems gewidmet. Als Erstes betrachten wir das Problem der verteilten Optimierung in einem vernetzen System. Mithilfe der Theorie der stochastischen Differenzialgleichungen analysieren wir eine stochastische Version der zeitkontinuierlichen verteilten Subgradienten Methode und leiten eine nicht-asymptotische hoch wahrschein-

liche Schranke sowie Erwartungsschranke für die Distanz zwischen dem Durchschnittsfunktionswert des Agenten Iterierten und dem optimalen Funktionswert her. Die Schranke liefert Empfehlungen von Parameterwahlen, die den Erfolg des Algorithmus garantieren. In dem Rest dieses Teils untersuchen wir das Verhalten der Meinungsdynamiken einer Population, die dem Einfluss von informationeller Voreingenommenheit unterliegt. Wir führen den neuen Begriff der substochastischen Komplementation ein. Dieser Begriff stellt einen effizienten Weg bereit, um das Netzwerk in Gruppen zu separieren. Außerdem erlaubt dieser uns den Grad der Gruppenartigkeit der Agentenmeinungen zu messen.

Acknowledgements

I am very grateful to my advisor Prof. Holger Boche for giving me his constant supports, a great deal of freedom in defining my research agenda, and finding my own interests. Without his encouragement, I would not be able to work on a wide range of topics, from control theory and convex optimization, to signal processing, machine learning, and statistics. I also thank Prof. Eduard A. Jorswieck for aggreeing to be the examiner of my thesis and Prof. Kellerer to be the chair of the examination commitee of my thesis.

I would also like to thank my colleagues at LTI, who made my time during my doctoral studies such a pleasure. Each of them was a constant source of inspiration and surely deserves a special mention. However, I limit myself to the essential ones. I want to thank Dr. Gisbert Janssen for the numerous discussions, including (but not limited to) the exciting topic of matrix analysis. It was also a pleasure to design and deliver the lecture "Matrix Concentration Inequalities and Applications" with you. I also want to thank Sebastian Baur and Alihan Kaplan for numerous discussions on various interesting topics. A special thanks go to Sajad Saeedinaeeni for being a terrific office mate.

I would like to thank the Surf WG Bali squad, especially Galih ijonk and Eddi Susilo, for their hospitality, and for the interesting mix of awesome surf and skate sessions and interesting chess games during the completion of this thesis in the corona-time in Bali.

I would thank all my friends, especially Christoph Klose and Jakob Jungmaier, for being there for me all the years, and for the numerous adventures.

I cannot thank my family: my parent Ebenezer Tampubolon and Tapi Ratna Tambunan and my brother Amos Tampubolon. I want to thank them for their support and advice throughout the years, for teaching me the importance of knowledge and learning, and for encouraging me to pursue my goals.

I am lucky to have Regine Hartwig in my life. First and foremost, I would thank her for reminding me that there is a life outside of work and the importance of work-life-balance. I am thankful to her for walking along my side with patience and support. I am also indebted to her for reading through my papers many times.

Contents

1. Introduction

Networked System in the Society and Technical Applications Large-scale interaction or relation in the form of a network between different entities has become a characteristic of several parts of our society. For instance, our economic market consists of enterprises (suppliers, focal company, customer) networked by flows of products, services, information, and financial means. Moreover, the mutual interference between the firms in the form of external economies/diseconomies (see, e.g., [1]) gives rise to another interaction-based structure. Another example is the relationship between political and social actors, which can be modeled employing a network, providing several advantages for the corresponding field of study (see, e.g. [2–4]. Also, the interaction-based view is indispensable for a vast number of groundbreaking present and future technical applications. A few examples are:

- The wireless sensor network offers low-cost efficient and robust solution for several important monitoring tasks such as military target tracking and surveillance, natural disaster relief, biomedical health monitoring, and seismic sensing (see, e.g., [5]).
- The network of cooperative aerial or ground robots provides solutions for several societal tasks, e.g., search and rescue operations in post-disaster environments, surveillance, factory automation, and logistics (see, e.g., [6–10])
- The smart grid, i.e., the power network composed of autonomous communicating intelligent nodes, is expected to improve the efficiency, reliability, and robustness of power or energy grids (see, e.g., [11, 12]).
- Wireless network users compete for base station utilization and generate in this way a network of Quality-of-Service externalities [13].
- The Internet-of-Things [14] shapes the interaction of a large number of heterogeneous and physically distributed sensing devices, communication technologies, and required services, enabling promising future concepts such as the concept of the smart city [15].

Modeling a Networked System as a Multi-Agent System – Cooperative vs. Non-Cooperative All the systems mentioned above have the similarity that they consist of a compound of subsystems or agents. Typically, those agents are autonomous in the sense that there is no direct intervention of superordinate bodies. Because if such instances exist, then surely the tasks they have to fulfill are of high complexity due to the typical number of subsystems. Moreover, by the scale reason, same as the latter, it is appropriate to assume that the agents are not fully aware of the state of the whole complex system. Besides this similarity, there is one principle difference between those multi-agent systems leading to their categorization into cooperative and non-cooperative multi-agent systems.

The main feature of a cooperative multi-agent system is that the interaction between agents is active, in the sense that they can communicate with each other. This feature allows the agents to interchange local information to infer global unknowns. For instance, in the wireless sensor network applications, each of the low-cost sensors can collect local data such as temperature, illuminance, or sound/level, to monitor the environment (see, e.g., [16]). The agents' communication feature can also be used to fulfill a global objective, such as done in the robotic network applications, where, e.g., drones swarm coordinate with each other to perform a search and rescue mission [17].

In contrary to the cooperative systems, the agents in a non-cooperative multi-agent system cannot interact directly via communication. So they are no able to self coordinate toward a mutual goal. The interaction between agents in such systems is on the benefit/cost level, in the sense that the action of an agent influences the benefit/cost structure of other agents. For example, in the wireless network applications where the users aim to optimize their transmission quality by adjusting the corresponding transmission power, the transmit power allocation of a user affects the transmission quality of other users because of transmission interference [13]. Another example of a non-cooperative multi-agent-system is the system of competitive firms in the economic market: The firms aim to optimize their earning by choosing their production output of a good while the production output of good affects the price of the good and consequently the earning of all the firms.

Machine-Learning Paradigm in Multi-Agent Systems Machine learning enables a system to deduce knowledge automatically. It designs a program/model for a given application that fits the data. This paradigm goes beyond the traditional programming paradigm, where programs are written to automate tasks specific to the corresponding application. For this reason, machine learning is capable of making techniques in widespread real-world applications efficient, flexible, resilient, and scalable. For instance, it has dramatically improved the field of medical imaging and computer-aided diagnosis, facilitating the healthcare sector (see, e.g., [18, 19]). Moreover, Search engines, which are indispensable tools for daily life, extensively use machine learning for tasks such as query suggestions, spell correction, web indexing, and page ranking [20,21]. Also, applications, where usually multi-agent systems occur, take advantage of machine learning. For instance, the field of communication looks forward to applying machine learning techniques for the design of the future communication system, which is essential for the progress of other applications such as vehicular technology [22]. The progress in the field of robotic is unimaginable without machine learning [23, 24]. Furthermore, as we look forward occupy more aspects of our lives with smart multi-agent technologies, ranging from home automation to autonomous vehicles, ML techniques will become increasingly crucial by aid the systems in decision making, analysis, and automation.

Challenges One of the main challenges related to multi-agent systems is to understand their behavior for the sake of eliminating the possible drawbacks occurring in their operation. One aspect toward this direction is to investigate the robustness of the system against possible error in the form of malicious agents' behavior (e.g., fake news) or imperfectness of the information obtained from the agents' environment (e.g., other agents or adversary nature). Another essential aspect is understanding the propagation of interaction within the population of agents, whether cooperative or non-cooperative. Another main challenge related to multi-agent systems is to design for a specific population-wide goal, such as efficiency and sustainability of a system, a control rule leading the agents toward the achievement of the goal. However, the control rule has to meet modern technologies' requirements, such as low complexity, scalability, privacy-preserving, and efficiency of information usage. Also, an important aspect is to involve present promising paradigm into consideration, such as the paradigm of machine learning.

Goal of this Thesis Our goal is to address above challenges with the following emphasize: In the first part of the thesis, we focus on robustness analysis and control mechanism design of non-cooperative learning agent systems; In the second part, on the design of an efficient algorithm for a single agent; and in the third part, on efficient algorithm design and structural analysis of cooperative systems.

1.1. Outline and Contributions

1.1.1. Part I: Resource Sustainable Robust Online Learning in Games

In the first part of this thesis, we study systems composed of non-cooperative strategic online-learning agents. Our aim is twofold. First, we want to understand the extent to which the feedback disturbance due to the lack of a global view is allowable, for that such a system is still stable. Second, we want to design an incentive control mechanism based on the state of the resources aiming to foster sustainable behavior in such systems of selfish learning agents. The outline and our contributions in this part of the thesis are explicitly stated in the following:

- In Chapter 3, we consider the class of games, called aggregative games, with the feature that the payoff of each player can be expressed to depend on his action and the aggregate of the population's action. This class of games serves as a model for a vast number of engineering fields, including signal processing and communications, such as communications, smart grid, and congestion control. We study the case, where the agents are unaware of the change of the aggregate, and where additionally they have only an estimate rather than the actual value of the aggregate. Our main contribution in this chapter is a mild sufficient condition on the learning agent's estimate of the aggregate depending on the step size, such that the population converges to the Wardrop equilibrium.

- Chapter 4 deals with general non-cooperative games, underlying coupled (resource) constraints. We assume that the players are learning agents who have merely noisy first-order utility feedback. This setting constitutes a model for widespread modern large-scale applications where the action of instances impacts the increase of scarce resources' congestion. Motivated by the aim to establish sustainable behavior by ensuring the fulfillment of the coupled constraints, and the aim to establish a stable population state, we propose a novel decentralized pricing mechanism via augmentation of the game's Lagrangian. For polynomially decaying learning agents' step size/learning rates, we provide a theoretical guarantee for achieving the latter aims by showing the almost sure convergence of the population's dynamic to the corresponding generalized Nash equilibrium. Also, investigate the finite-time quality of the proposed algorithm by giving a non-asymptotic time decaying bound for the amount of resource constraint violation and the distance between the population's state and the generalized equilibrium.

- In Chapter 5, we extend our contributions in Chapter 4 by introducing an additional parameter in the proposed price mechanism reflecting the agents' price sensitivity. We show theoretically that, to a particular extent, the agents' sensitization for the prices results in sustainable behavior. We show this by providing a sub-linear bound for the aggregate of the constraint violations. Furthermore, we show numerically that over sensitization of the agents for prices results in a population's behavior contrary to resource sustainability.

- In Chapter 6, we turn our attention to the class of congestion games, which constitutes a model for many resource allocation problems, such as network routing problems and wireless channel allocation problems. Based on the assumption of rational non-cooperative cost-oriented (and not feedback oriented) agents, we propose a novel resource-centric dynamic pricing, that offers the system participants appropriate incentives to adhere to the resource constraints. We show theoretically and numerically that the proposed pricing mechanism ensures the sub-linear decay of the average violation of the capacity constraints. Also, we show that, although it does not use specific information about the agents, our pricing mechanism does not significantly affect the agents' welfare and may even result in the latter's improvement.

1.1.2. Part II Online Decision-Making with Emphasize on the Noise-Robustness and Sustainable Behaviour

The second part of this thesis is devoted to the microscopic aspect of the multi-agent systems and the study of online learning agent problems with constraints. Driven by the drawback of state of the art algorithms, concentrating on establishing sub-linear growth of the merely cumulative long-term constraint violations, we introduce a new tighter performance measure called h-CFit. We propose a class of non-causal algorithms for online-decision making, which guarantees, in slowly changing environments, sub-linear growth of this quantity despite noisy first-order feedback. We demonstrate by numerical experiments the performance gain of our method relative to state of art.

1.1.3. Part III: Distributed Coordination Algorithms

In contrast to the first part of this thesis, we study in the third part of this thesis, locally cooperative multi-agent system where neighboring agents can exchange information. Our main aim is to understand how perturbations in the form of noise and extrinsic opinion influence, such as fake news, affect such systems' global behavior. The specific outline and the contributions of this part are as follows:

- In Chapter 8, we consider the problem of distributed optimization. For finding the solution to this problem, we propose a novel variant of continuous-time distributed gradient descent algorithm, where the gradient is subject to Gaussian noise contamination. Our object of study is the influence of the underlying model parameters, i.e., the function's parameters and the connectivity of the agents, the parameters of the dynamic, i.e., the step size/gradient weight and the communication strength between the agents, and the volatility of the noise process, to the success of the algorithm. Our main contribution is a bound quantifying the decay of the distance between the agents' objective value and the consensus optimum, both in expectation and high probability.

- In Chapter 9, we study the impact of extrinsic influence in the form of informational bias, e.g., fake news, to the population's opinion. Our main aim is to formally show that informational bias results in the establishment of mesoscopic stability, meaning that the population's opinion is cluster-dispersive. Toward this direction, we propose the novel notion of substochastic complementation, which provides an efficient way to approximate the population's dynamic by cluster dynamics. Motivated by this notion, we propose a novel measure for cluster-dispersion of opinion dynamic in the face of the informational bias and analyze it for several limit cases of disturbances by informational bias.

1.2. List of Publications

1.2.1. Part I: Resource Sustainable Robust Online Learning in Games

Chapter 3 is based on the following publication:

- **E. Tampubolon** and H. Boche, "On the Convergence of Online Mirror Ascent for Aggregative Games with Approximated Aggregates," *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2019.

Chapter 4 is based on the following publication:

- **E. Tampubolon** and H. Boche, "Semi-Decentralized Coordinated Online Learning for Continuous Games with Coupled Constraints via Augmented Lagrangian," arXiv preprint arXiv:1910.09276, 2019.
- **E. Tampubolon** and H. Boche, "Coordinated Online Learning for Multi-Agent Systems with Coupled Constraints and Perturbed Utility Observations," *IEEE Transactions on Automatic Control* vol. 65, no. 11, 2020.

Chapter 5 is based on the following publications:

- **E. Tampubolon** and H. Boche, "Pricing Mechanism for Resource Sustainability in Competitive Online Learning Multi-Agent Systems," arXiv preprint arXiv:1910.09314, 2019.
- **E. Tampubolon** and H. Boche, "Robust Pricing Mechanism for Resource Sustainability under Privacy Constraint in Competitive Online Learning Multi-Agent Systems," in *45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020.

Chapter 6 is based on the following publications:

- **E. Tampubolon** and H. Ceribasic and H. Boche, "Resource-Aware Control via Dynamic Pricing for Congestion Game with Finite-Time Guarantees," arXiv preprint arXiv:2002.06080, 2020.
- **E. Tampubolon** and H. Ceribasic and H. Boche, "Resource-Aware Control via Dynamic Pricing for Congestion Game with Finite-Time Guarantees," in *21st IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020.

1.2.2. Part II Online Decision-Making with Emphasize on the Noise-Robustness and Sustainable Behaviour

The second part of this thesis is based on the following publications:

- **E. Tampubolon** and H. Boche, "Robust Online Learning for Resource Allocation - Beyond Euclidean Projection and Dynamic Fit," arXiv preprint arXiv:2002.06080, 2019.
- **E. Tampubolon** and H. Boche, "Robust Online Mirror Saddle-Point Method for Constrained Resource Allocation," in *45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2020.

1.2.3. Other Publications

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

Foundation of Signal Processing

- H. Boche and **E. Tampubolon**, "On the decay – and the smoothness behavior of the Fourier transform, and the construction of signals having strong divergent Shannon sampling series," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016
- H. Boche and **E. Tampubolon**, "On the Existence of the Band-Limited Interpolation of Non-Band-Limited Signals," *24th European Signal Processing Conference (EUSIPCO)*, 2016, 428-432

Sampling Method for Signals

- **E. Tampubolon**, V. Pohl, and H. Boche, "A Class of Frames for Paley-Wiener spaces with Multiple Lattice Tiles Support," *11th International Conference on Sampling Theory and Applications (SampTA)*, 2015.
- H. Boche, U.J. Mönich, **E. Tampubolon**, "Strong divergence of the Shannon sampling series for an infinite dimensional signal space," *IEEE International Symposium on Information Theory*, 2016, 2878-2882.
- H. Boche, U.J. Mönich, **E. Tampubolon**, "Structure of the Set of Signals with Strong Divergence of the Shannon Sampling Series," in *42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.
- H. Boche, U.J. Mönich, **E. Tampubolon**, "Spaceability and Strong Divergence of the Shannon Sampling Series and Applications," *Journal of Approximation Theory*, vol. 222, pp. 157-174, 2017
- **E. Tampubolon** and H. Boche, "On the Equivalence Between Classical and Distributional Convergence for Shannon Type Interpolation Series and Applications," *11th International ITG Conference on Systems, Communications and Coding*, 2017.

Peak-to-Average Power Reduction in Multicarrier Systems

- H. Boche and **E. Tampubolon**, "Mathematics of signal design for communication systems," In: König, W., ed. (Ed.): *Mathematics and Society*, European Mathematical Society Publishing House, pp. 185-220, 2016.
- H. Boche, U.J. Mönich and **E. Tampubolon**, "Complete Characterization of the Solvability of PAPR Reduction for OFDM by Tone Reservation," in *2017 IEEE International Symposium on Information Theory*, 2017.
- H. Boche and **E. Tampubolon**, "Asymptotic analysis of tone reservation method for the PAPR reduction of CDMA systems," in *IEEE International Symposium on Information Theory*, 2017.
- **E. Tampubolon** and H. Boche, "Probabilistic Analysis of Tone Reservation Method for the PAPR Reduction of OFDM Systems," in *42nd IEEE International Conference on Acoustics, Speech and Signal Processing*, 2017
- H. Boche and **E. Tampubolon**, "PAPR Problem for Walsh Systems and Related Problems," *IEEE Transactions on Information Theory* vol. 64, no. 8, pp. 5531-5548, 2018.

1.3. Basic Notations and Notions

Numbers, Vectors, and Matrices For $n \in \mathbb{N}_0$, $[n]$ (resp. $[n]_0$) denotes the set of integers between 1 (resp. 0) and n. We use boldface to distinguish between vectors (or sequence) and scalars and also between vector-valued and scalar valued functions. Upright letters stand for functions and matrices. We write random variable by capital letter, but capital upright letter stands for matrices. Given a vector/sequence $\boldsymbol{x}$. $\boldsymbol{x}_r$ denotes the r-th member of $\boldsymbol{x}$. We use the same notation for vector-valued function and random vector/sequence. In case we have sequence of vectors $\boldsymbol{\Lambda}$, we denote $\boldsymbol{\Lambda}_n^r$ as the r-th entry of the n-th member. Inequalities with matrices is meant entrywise. We denote the vector on $\mathbb{R}^D$ with all entries equal to 1 by $\mathbf{1}_D$. We denote the $D \times D$ identity matrix by $\mathbf{I}_D$. If the underlying dimension of the Euclidean space is clear, we sometimes write more compactly $\mathbf{1}$ and $\mathbf{I}$ for $\mathbf{1}_D$ and $\mathbf{I}_D$.

Normed Space In this work, we always consider the usual Euclidean space $\mathbb{R}^D$. For a convex subset $\mathcal{A} \subseteq \mathbb{R}^D$, $\mathrm{relint}(A)$ denotes the relative interior of $\mathcal{A}$, and if not otherwise stated, $\|A\|$ denotes the supremum of $\|\boldsymbol{x} - \boldsymbol{y}\|$ over all $\boldsymbol{x}, \boldsymbol{y} \in A$ called the diameter of $\mathcal{A}$ w.r.t. a norm $\|\cdot\|$ on $\mathbb{R}^D$. The projection onto a closed convex subset $\mathcal{A}$ of $\mathbb{R}^D$ is denoted by $\mathbf{\Pi}_{\mathcal{A}}$. The dual norm of a norm $\|\cdot\|$ on $\mathbb{R}^D$ is denoted by $\|\cdot\|_*$. $\mathcal{V}^*$ denotes the dual space of the Euclidean normed space $\mathcal{V} = (\mathbb{R}^D, \|\cdot\|)$ which is $\mathcal{V}^* = (\mathbb{R}^D, \|\cdot\|_*)$.

$\mathbf{g} : \mathbb{R}^D \to \mathbb{R}^D$ is said to be Lipschitz continuous on an a non-empty subset $\mathcal{Z} \subset (\mathbb{R}^D, \|\cdot\|)$ with constant $L > 0$ if $\|\mathbf{g}(\boldsymbol{x}) - \mathbf{g}(\boldsymbol{z})\|_* \leqslant L\|\boldsymbol{x} - \boldsymbol{z}\|$, $\forall \boldsymbol{x}, \boldsymbol{z} \in \mathcal{Z}$. F is said to be monotone on $\mathcal{Z}$ if $\langle \boldsymbol{x}_1 - \boldsymbol{x}_2, \mathbf{g}(\boldsymbol{x}_1) - \mathbf{g}(\boldsymbol{x}_2)\rangle \leqslant 0$, for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{Z}$. If in the latter strict inequality hold for $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, then $\mathbf{g}$ is said to be strictly monotone. $\mathbf{g}$ is said to be c-strongly monotone on $\mathcal{Z}$ if $\langle \boldsymbol{x}_1 - x_2, \mathbf{g}(\boldsymbol{x}_1) - \mathbf{g}(\boldsymbol{x}_2)\rangle \leqslant -c\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2$, for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{Z}$.

Landau Notation Throughout this thesis, we often utilize the Landau asymptotic notations such as $\mathcal{O}$, o, Ω, Θ. To define those notations, let $\mathrm{f}, \mathrm{g} : \mathbb{R} \to \mathbb{R}$. We say that f grows no faster than g asymptotically and write $\mathrm{f}(t) = \mathcal{O}(\mathrm{g}(t))$ (as $t \to \infty$) , if:

$$\limsup_{t\to\infty} \frac{\mathrm{f}(t)}{\mathrm{g}(t)} < \infty,$$

i.e., there exists some positive constant $c > 0$ such that $|\mathrm{f}(t)| < \mathrm{g}(t)$ for sufficiently large t. Conversely, we say that f grows no slower than g asymptotically and write $\mathrm{f}(t) = \Omega(\mathrm{g}(t))$ if:

$$\mathrm{g}(t) = \mathcal{O}(\mathrm{f}(t)),$$

or equivalently, if:

$$\limsup_{t\to\infty} \frac{\mathrm{g}(t)}{\mathrm{f}(t)} < \infty,$$

If we have both:

$$\mathrm{f}(t) = \mathcal{O}(\mathrm{g}(t)) \quad \text{and} \quad \mathrm{g}(t) = \mathcal{O}(\mathrm{f}(t)),$$

we say that f grows as g we write $\mathrm{f}(t) = \Theta(\mathrm{g}(t))$. Finally, if:

$$\limsup_{t\to\infty} \frac{\mathrm{f}(t)}{\mathrm{g}(t)} \leqslant 0,$$

we write $\mathrm{f}(t) = o(\mathrm{g}(t))$ and we say that f is dominated by g. We sometimes also use the Landau notation for Matrix-valued function, such as $\mathbf{A}(t) = \mathcal{O}(\mathrm{f}(t))$, where $\mathbf{A} : \mathbb{R} \to \mathbb{R}^{D_1 \times D_2}$ and $\mathrm{f} : \mathbb{R} \to \mathbb{R}$. This means that $\|\mathbf{A}(t)\| = \mathcal{O}(\mathrm{f}(t))$, where $\|\cdot\|$ denotes any matrix norm. The same also holds for other Landau notations.

Part I.

Resource Sustainable Robust Online Learning in Games

2. Preliminaries

2.1. Online Learning

2.1.1. Generic Setting

The online learning [25] is an emerging paradigm aiming to solve the problem of sequential decision making in an unknown and possibly adversarial environment. The setting is as follows (see also Algorithm 2.1.1): Consider a time horizon $T \in \mathbb{N}$. At each time slot $t \in [T]$, the decision-maker takes an action $\boldsymbol{x}_t$ from a prespecified set $\mathcal{X}$. The environment responds to the action $\boldsymbol{x}_t$ of the decision-maker and charges the decision-maker the cost $\mathrm{f}_t(\boldsymbol{x}_t)$, where f_t is an apriori unknown loss function. Subsequently, the decision-maker selects a new action $\boldsymbol{x}_{t+1}$ for the next stage $t+1$. In order to do that she might query information about the present and/or the historical cost functions, such as the first-order information, i.e., the gradient of the functions. We refer the way how the decision-maker selects the future action in this context as *online learning policy*.

Algorithm 1 Generic Online Learning Process

Require: Action set $\mathcal{X}$, time horizon $T \in \mathbb{N}$, loss/cost function $\mathrm{f}_t : \mathcal{X} \to \mathbb{R}$.
  for $t = 0, 1, 2, \ldots, T$ **do**
    Take the action $\boldsymbol{x}_t \in \mathcal{X}$
    Incur cost/loss $\mathrm{f}_t(\boldsymbol{x}_t)$
    Update:

$$(\boldsymbol{x}_t,\ \text{information about } (\mathrm{f}_\tau)_{\tau \in [t]_0}) \mapsto \boldsymbol{x}_{t+1} \tag{2.1}$$

  end for

The online learning method is ideally suited for application where the underlying problem is subject to unpredictable dynamic, such as dispatch of renewable energy having intermittent and unpredictable nature, or network applications where the task to accomplish is subject to unpredictable human participation, or applications requiring flexibility in handling heterogenity and scalability. Furthermore, applications requiring real-time decision leverage from an online learning method since the given algorithm is usually lightweight. For those reasons, online learning has become in the recent years a popular method to solve several resource allocation and management problems in several engineering fields such as economic dispatch in power systems [26, 27], data center scheduling [28–30], electric vehicle charging [31, 32], video streaming [33], thermal control [34], and fog computing in IoT [35–37].

In online learning, the aim of the decision-maker is to minimize the so-called (cumulative) regret defined as:

$$\mathrm{R}_T = \mathrm{R}_T((\boldsymbol{x}_\tau)_{t\in[T]}) = \sum_{t=0}^{T} \mathrm{f}_t(\boldsymbol{x}_t) - \inf_{\boldsymbol{x}\in\mathcal{X}} \sum_{t=0}^{T} \mathrm{f}_t(\boldsymbol{x}).$$

In words the cumulative regret measures the sub-optimality of the cumulative loss of the decision-maker with respect to the cumulative loss of the best action in hindsight. More optimistic aim than regret minimization such as minimization of the present cost function is in general in feasible, since in order to decide for the future action the decision-maker only knows about the present cost function. One direction to achieve the aim of regret minimization is by means of the so-called no-regret policy (see e.g. [38–40]). The *no-regret policy* means in this context that during each round of the game, the agent endeavors to choose an action such that the *regret*, i.e., the cumulative difference between the instantaneous yields and the corresponding highest possible yields, grows slower than the number of rounds and decays if averaged to zero as the game progresses. Specifically, the no-regret policy comprises for any $T \in \mathbb{N}$ methods to choose actions $(\boldsymbol{x}_t)_{t\in[T]}$ according to the procedure described in Algorithm 1 satisfying:

$$\mathrm{R}_T = o(T).$$

In order to illustrate the concept of online learning, we provide in the following soma examples:

Example 1 (Prediction with Expert Advice): The classical example of online paradigm is the so-called prediction with expert advice. The setting is as follows. On each round t, the learner chooses one advice given by $N > 0$ experts. Thus $\boldsymbol{x}_t \in \{\mathbf{e}_i\}_{i=1}^N \subseteq \mathbb{R}^N$ where $\{\mathbf{e}_i\}_{i=1}^N$ denotes the usual orthonormal basis of $\mathbb{R}^D$, and where for every $i \in [D]$:

$$\boldsymbol{x}_t^i = \begin{cases} 1, & \text{if the learner chooses the advice of expert } i, \\ 0, & \text{Otherwise} \end{cases}.$$

Taking the advice of expert i, the learner incurs loss $\boldsymbol{y}^(i)_t$ and observes each losses obtained by taking the advice of the experts. The latter can be described by the vector $\boldsymbol{y}_t \in \mathbb{R}^D$. We can finally adapt this procedure to Algorithm 1 by defining the cost function as follows:

$$\mathrm{f}_t(\boldsymbol{x}) = \langle \boldsymbol{y}_t, \boldsymbol{x} \rangle, \quad \boldsymbol{x} \in \mathbb{R}^D$$

Example 2 (Prediction with Experts - Convexification via Randomization): Working with the online setting of the prediction with expert advice (see Example Example 1) is cumbersome since the action set is discrete and therefore the corresponding problem is of combinatorial nature. In order to leverage from the mathematical analysis, we can modify the setting given in Example 1 as follows:

As the action set of the learner we take instead the simplex Δ on $\mathbb{R}^D$, where for an action $\boldsymbol{x}_t \in \mathcal{X} = \Delta$, $\boldsymbol{x}_t^i$ denotes the probability that the learner i follows the advice of expert i. After taking the action $\boldsymbol{x}_t$ on the round t, the learner observes the cost $\boldsymbol{y}_t \in \mathbb{R}^D$ and incurs the corresponding loss measured by the expectation:

$$\mathrm{f}_t(\boldsymbol{x}_t) = \langle \boldsymbol{y}_t, \boldsymbol{x}_t \rangle.$$

In this setting both the action set and the loss functions are convex. Therefore, we can leverage from the field of convex optimization having extensive literature.

Example 3: The next example comes from the signal processing application and concerns with the problem of signal covariance optimization in multiple-input and multiple-output (MIMO) wireless networks (see e.g. [41]). Here, we consider a multi-user MIMO wireless network where a set of autonomous wireless devices $[N]$ each equipped with multiple antennas seek to maximize their individual data rates. For a device $k \in [N]$, its Shannon achievable rate is given by:

$$C(\mathbf{Q}_k, \tilde{\mathbf{H}}_k) = \log\det(\mathbf{I} + \tilde{\mathbf{H}}_k \mathbf{Q}_k \tilde{\mathbf{H}}_k^{\mathrm{T}}),$$

where $\mathbf{Q}_k$ denotes the signal covariance matrix of device k, and $\tilde{\mathbf{H}}_k$ denotes its effective gain channel matrix (see e.g. [41] or [42]). In a realistic setting, $\tilde{\mathbf{H}}_k$ changes over time. Moreover, $\tilde{\mathbf{H}}_k$ encompasses in particular the effects of the wireless channel such as noise, path loss, and path diversity and depends on the transmit characteristics of all interfering users. Therefore generally, the effective channel gain matrix cannot be known in advance. This occurence suggest to tackle the problem of signal covariance optimization in MIMO wireless networks by online learning paradigm. For this sake, one can define the action set of the learner, i.e. the device k, as:

$$\mathcal{X} = \{\mathbf{Q} : \ \mathrm{tr}(\mathbf{Q}) \leqslant P_{\max}\},$$

where $P_{\max} > 0$ denotes the user's maximum transmit power. Finally, one can set the loss function of the learner by:

$$\mathrm{f}_t(\mathbf{Q}_t) = -C(\mathbf{Q}_t, \mathbf{H}_t).$$

2.1.2. Online Gradient Descent

The straightforward approach for solving optimization problems in case that the first-order information of the objective function is available is by the so-called (projected) gradient descent method, where at each stage one takes a step in direction of the steepest descent of the objective function, i.e. the gradient of the function, and if necessary, projects the

Table 2.1.: Regret bound under L-lipschitz assumption

Assumptions	Minimax Regret	Worst case Bound
Convex loss function	$\Omega(L\sqrt{T})$	$\mathcal{O}(L\sqrt{T})$
Linear loss function	$\Omega(L\sqrt{T})$	$\mathcal{O}(L\sqrt{T})$
β-strongly convex	$\Omega(\beta^{-1}L^2 \log T)$	$\mathcal{O}(\beta^{-1}L^2 \log T)$

iterate to the problem's feasible region. The seminal work [25] extends this method to the online learning setting by defining the update as:

$$\boldsymbol{x}_{n+1} = \boldsymbol{\Pi}_{\mathcal{X}}\left(\boldsymbol{x}_n - \gamma_n \mathbf{v}_n\right), \tag{2.2}$$

where γ_n denotes the step size at stage n and where:

$$\mathbf{v}_n := \nabla_{\boldsymbol{x}_n} \mathrm{f}_n(\boldsymbol{x}_n)$$

Under certain condition on the underlying setting, the online gradient descent satisfies the no-regret property. One basic example of such guarantee is the following:

Theorem 2.1: *Suppose that $\mathcal{X}$ is convex and non-empty, and for any $n \in \mathbb{N}$, f_n is convex and L-Lipschitz. For a time horizon $T \in \mathbb{N}$, it holds:*

$$\mathrm{R}_T \leqslant L\|\mathcal{X}\|_2\sqrt{T},$$

if the step-size is chosen to be constant with:

$$\gamma_n = \frac{\|\mathcal{X}\|_2}{LT}$$

So according to above Theorem, provided that the function is convex and Lipschitz, the regret at the end of a time horizon increases slower than the length of the considered time interval. By this reason one refers such bound also as *worst-case bound.* Further assumption on the underlying loss leads to slower increase of the regret (see Table 2.1).

In general, the worst-case bound given in Theorem 2.1 cannot be improved (up to a constant), since one can construct (see [43]) a sequence of loss functions, such that the regret is not less than $\mathcal{O}(\sqrt{T})$:

$$\mathrm{R}_T \geqslant \frac{L\|\mathcal{X}\|_2}{2\sqrt{2}}\sqrt{T}.$$

The sort of bound given above is also called *minimax bound* because it provides an estimate of the minimum value for a worst-case scenario. Another bounds for further subclasses of loss function is provided in Table 2.1

Even though the order of the worst-case regret guarantee of OGD is tight, there is still some possibilities to improve it. For instance, the regret bound given in Theorem 2.1 depends on quantities, i.e. the diameter $\|\mathcal{X}\|_2$ and the Lipschitz constant L, measured by means of the Euclidean norm, which is known to suffer from the curse of dimensionality. In the modern application (such as big data applications) where the underlying problem space has typically tremendeously high dimension such drawback might cause serious problems.

2.1.3. Online Mirror Descent

One solution to avoid the curse of dimensionality when applying the descent method is to move one's cosideration beyond the classical Euclidean geometry. A systematic way to exploit the geometry of the problem is by generalizing the projection step in the update of the OGD. One possible way is by the so called mirror map defined as follows:

Definition 2.1 (Regularizer/Penalty and Mirror Map): *Let $\mathcal{Z}$ be a compact convex subset of a normed space $(\mathcal{E}, \|\cdot\|)$, and $K > 0$.*

- *We say $\psi : \mathcal{Z} \to \mathbb{R}$ is a K-strongly convex regularizer (or penalty function) on $\mathcal{Z}$, if ψ is continuous and K-strongly convex on $\mathcal{Z}$.*
- *The mirror map $\mathbf{\Phi} : \mathcal{E}^* \to \mathcal{Z}$ induced by ψ is defined by:*

$$\mathbf{\Phi}(\boldsymbol{y}) := \arg\max_{\boldsymbol{x}\in\mathcal{Z}} \{\langle \boldsymbol{y}, \boldsymbol{x}\rangle - \psi(\boldsymbol{x})\}$$

.

The following example shows that the mirror map is indeed a generalization of the usual Euclidean projection:

Example 4 (Euclidean projection): Let $\mathcal{Z}$ be compact convex subspace of a Euclidean space. Clearly,

$$\psi(\boldsymbol{x}) = \frac{\|\boldsymbol{x}\|^2}{2}$$

is a 1-strongly convex regularizer on $\mathcal{Z}$. Short computation yields that the induced mirror map is the Euclidean projection onto $\mathcal{Z}$, i.e.:

$$\mathbf{\Phi}(\boldsymbol{y}) = \arg\max_{\boldsymbol{x}\in\mathcal{Z}} \|\boldsymbol{y} - \boldsymbol{x}\|_2.$$

Another interesting example of mirror maps which is also popular in the field of decision making is the following:

Example 5 (Entropic regularization & Logit choice): The so called logit choice:

$$\mathbf{\Phi}(\boldsymbol{y}) = \frac{\exp(\boldsymbol{y})}{\sum_{l=1}^{D} \exp(\boldsymbol{y}_l)}$$

is generated by the 1-strongly convex regularizer

$$\psi(\boldsymbol{x}) = \sum_{k=1}^{D} \boldsymbol{x}_k \log \boldsymbol{x}_k,$$

known as the Gibbs entropy or the negative Shannon's entropy, on the probability simplex:

$$\mathcal{Z} = \Delta \subset (\mathbb{R}^D, \|\cdot\|_1).$$

Notice that in contrast to the Euclidean projection on the simplex (see e.g. [44]) the logit choice has a closed form and therefore easier to implement.

Another non-standard examples of mirror maps are the following:

Example 6 (Fermi-Dirac Entropy): Suppose that:

$$\mathcal{Z} = [0,1]^D$$

is the unit cube. The Fermi-Dirac entropy:

$$\psi(\boldsymbol{x}) = \sum_{k=1}^{D} (\boldsymbol{x}_k \log \boldsymbol{x}_k + (1-\boldsymbol{x}_k)\log(1-\boldsymbol{x}_k).$$

induces the mirror map:

$$\boldsymbol{\Phi}(\boldsymbol{y}) = \left(\frac{\exp(\boldsymbol{y}_k)}{1+\exp(\boldsymbol{y}_k)}\right)_{k\in[D]}$$

called logistic map.

Example 7 (Matrix regularization): Let $\mathcal{Z}$ be the set of positive semidefinite matrices $\mathbf{X}$ having the nuclear norm:

$$\|\mathbf{X}\|_1 := \mathrm{tr}(|\mathbf{X}|) \leqslant 1.$$

We consider $\mathcal{Z}$ as the subspace of the Euclidean space of symmetric matrices. The von-Neumann entropy:

$$\psi(\mathbf{X}) = \mathrm{tr}(\mathbf{X}\log\mathbf{X}) + (1-\mathrm{tr}\mathbf{X})\log(1-\mathrm{tr}\mathbf{X})$$

is a (1/2)-strongly convex regularizer on $\mathcal{Z}$ [45]. It induces the mirror map (for derivation see e.g. [46]):

$$\boldsymbol{\Phi}(\mathbf{Y}) = \exp(\mathbf{Y})/(1+\|\exp(\mathbf{Y})\|_1).$$

Having introduced the notion of mirror map, we can generalize the online gradient descent (2.2)as follows:

$$\begin{aligned} \boldsymbol{y}_{t+1} &= \boldsymbol{y}_t - \gamma_t \mathbf{v}_t \\ \boldsymbol{x}_{t+1} &= \boldsymbol{\Phi}(\boldsymbol{y}_{t+1}). \end{aligned}$$

Above method is known as online mirror ascent, which is a canonical extension of the mirror ascent algorithm [47]. Besides providing a richer model for the online decision-making process by generalizing projected gradient descent, a well-known advantage of mirror descent is that by appropriate choice of the mirror step the performance of the corresponding algorithm might have a weaker influence on the dimension of the underlying decision space (see e.g., [48]).

One of the basic worst-case guarantee for online mirror descent is the following:

Theorem 2.2: *Suppose that the losses are L-Lipschitz. The online mirror descent induced by a K-strongly convex regularizer satisfies:*

$$\mathrm{R}_T \leqslant 2L\sqrt{\frac{\max_{\mathcal{X}} \psi(\boldsymbol{x}) - \min_{\mathcal{X}} \psi(\boldsymbol{x})}{2K}T}, \tag{2.3}$$

in case that the step size is chosen by:

$$\gamma_t = \frac{1}{L}\sqrt{\frac{2K(\max_{\mathcal{X}} \psi(\boldsymbol{x}) - \min_{\mathcal{X}} \psi(\boldsymbol{x}))}{T}}$$

Notice that in the above guarantee, the strong convexity - and the Lipschitz constants K and L need not be taken with respect to the Euclidean norm which can be crucial as shown in the following:

Example 8: Suppose that f is differentiable and is 1-Lipschitz w.r.t. $\|\cdot\|_1$, i.e.:

$$\|\nabla \mathrm{f}(\boldsymbol{x})\|_\infty \leqslant 1, \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Therefore, the inequality $\|\cdot\|_2 \leqslant D\|\cdot\|_\infty$ implies that f is D-Lipschitz w.r.t. $\|\cdot\|_\infty$. Thus by using Euclidean norm instead of the maximum norm, we obtain additional dimensional factor in our guarantee.

Now, we discuss in the following above result for several choices of mirror maps:

Example 9: Online Gradient Descent As we have discussed in Example 4, online gradient descent is simply the online mirror descent with the Euclidean norm (divided by 2) as the 1-strongly convex regularizer. Since:

$$\Delta\psi(\mathcal{X}) = \max_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{x}\|_2 - \min_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{x}\|_2 \leqslant \|\mathcal{X}\|_2,$$

we have from (2.2) the worst-case guarantee of order $\mathcal{O}(L\sqrt{T})$ coinciding with the worst-case guarantee given in Theorem 2.1 up to a multiplicative constant 2. This gap can be eliminated by using the "centered" regularizer:

$$\psi(\boldsymbol{x}) = \frac{\|\boldsymbol{x} - \boldsymbol{x}_c\|_2}{2},$$

where $\boldsymbol{x}_c$ is the center of the smallest Euclidean ball enclosing $\mathcal{X}$.

Example 10 (Exponentially Weights Algorithm): Suppose that the underlying feasible set is the simplex, and that f is differentiable and 1-Lipschitz w.r.t. $\|\cdot\|_\infty$. Choosing the Gibbs entropy as the 1-convex regularizer w.r.t. the $\|\cdot\|_1$ and the logit choice as the mirror map, the online mirror descent is also known as the exponentially weights algorithm. It holds:

$$\Delta\psi = \log(D),$$

leading to the worst-case regret bound (2.3) of order $\mathcal{O}(\sqrt{\log(D)T})$. Using instead the Euclidean norm $\psi(\boldsymbol{x}) = \|x\|_2/2$, it holds $\Delta\psi = \|\Delta\|_2 = \sqrt{2}$. Moreover, Example 8 asserts that $L = D$. Thus we have the the worst-case regret bound (2.3) of order $\mathcal{O}(\sqrt{\log(D)T})$ which is worst for high-dimensional problem.

The following Proposition which is a folklore in convex analysis gives some basic properties of the mirror map:

Proposition 2.3: *Let ψ be a K-strongly convex regularizer on a compact convex subset $\mathcal{Z}$ of a Euclidean normed space $\mathcal{V}$ inducing the mirror map $\boldsymbol{\Phi} : \mathcal{V}^* \to \mathcal{Z}$, and let $\psi^* : \mathcal{V}^* \to \mathbb{R}$, $y \mapsto \max_{\boldsymbol{x}\in\mathcal{Z}} \{\langle x, y\rangle - \psi(x)\}$ be the convex conjugate of ψ. Then:*

1. *$x = \boldsymbol{\Phi}(y)$ if and only if $y \in \partial\psi(x)$. In particular im$(\boldsymbol{\Phi})$ = dom$(\partial\psi) \supseteq$ relint$(\mathcal{Z})$.*

2. *ψ^* is differentiable on $\mathcal{V}^*$ and $\nabla\psi^*(y) = \boldsymbol{\Phi}(y)$.*

3. *$\boldsymbol{\Phi}$ is $(1/K)$-Lipschitz continuous.*

4. *ψ is $1/\|\mathcal{Z}\|_*$-strongly convex w.r.t. $\|\cdot\|$.*

Proof: For a proof of 1)-3), see e.g. Theorem 23.5 in [49] and Theorem 12.60(b) in [50].

For the statement 4), notice that $\|\nabla\psi^*(y)\|_* = \|\boldsymbol{\Phi}(y)\|_* \leqslant \|\mathcal{Z}\|_*$ where the inequality follows from the fact that $\boldsymbol{\Phi}$ is a mapping to $\mathcal{Z}$. Therefore ψ^* is $\|\mathcal{Z}\|_*$-strongly smooth and Strong/smooth duality Theorem (see e.g. Theorem 3 in [45]) asserts the desired statement. ■

As noticed in [51], a convex regularizer induces canonically the following notion of "distance":

Definition 2.2 (Fenchel Coupling [51]): *Let $\psi : \mathcal{X} \to \mathbb{R}$ be a penalty function on $\mathcal{X}$. Then the Fenchel coupling induced by ψ is defined as:*

$$\mathrm{F} : \mathcal{X} \times E^* \to \mathbb{R}, \quad (p, y) \mapsto \psi(p) + \psi^*(y) - \langle y, p\rangle.$$

Some useful properties of the Fenchel coupling is stated in the following (see [51]):

Proposition 2.4: *Let* F : *be the Fenchel coupling induced by a* K*-strongly convex regularizer of* $\mathcal{X}$*. For* $p \in \mathcal{X}$, $y, y' \in \mathcal{V}^*$, *we have:*

1. $\mathrm{F} = (p, y) \geqslant \frac{K}{2}\|\mathbf{\Phi}(y) - p\|^2$

2. $\mathrm{F} = (p, y') \leqslant \mathrm{F}(p, y) + \langle y' - y, \mathbf{\Phi}(y) - p\rangle + \frac{1}{2K}\|y' - y\|_*^2$

For a proof of those facts, see e.g. Theorem 23.5 in [49] and Theorem 12.60(b) in [50].

2.2. Game Theory

The concept of *game theory* is created to describe, analyze, and forecast system behavior in such population models with information exchange constraints. It considers a set of agents whose aim is, for rationality assumption, to maximize their own yield/payoff or equivalently to minimize their loss by choosing appropriate actions. The feature of the game-theoretical concept is that an agent's yield depends not only on her action but also on the action of others. Furthermore, game theory assumes that the agents are non-cooperative in the sense that the strategy of one agent is not visible to the others. One of the main goals of the game theory is to predict, in a repeated setting and for an individual specific rational optimizing behavior of the agents, the long-term behavior of the population [52,53]. The central sub-concept toward this direction is the so-called *Nash equilibrium* [54], which denotes a stable state where no agent has an interest in deviating from her strategy.

2.2.1. Basic Definitions

We consider throughout this part of thesis agents playing a (repeated). *non-cooperative game.* During the non-cooperative game, every agent $i \in [N]$ chooses and applies an action/strategy $\boldsymbol{x}^{(i)}$ from a set $\mathcal{X}_i$. This process results in joint action/strategy-profile:

$$\boldsymbol{x} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}) \in \mathcal{X} := \prod_{i=1}^{N} \mathcal{X}_i \in \mathbb{R}^D, \quad \text{where} \quad D := \sum_{i=1}^{N} D_i.$$

In order to highlight the action of player i, we write:

$$\boldsymbol{x} = (\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(-i)}), \quad \text{where} \quad \boldsymbol{x}^{(-i)} = (\boldsymbol{x}^{(j)})_{j \neq i} \in \mathcal{X}_{-i} := \prod_{j \neq i} \mathcal{X}_j.$$

Unless otherwise stated, we assume throughout this thesis the following:

Assumption 2.1: *For all* $i \in [N]$, $\mathcal{X}_i$ *is a non-empty compact convex subset of a finite dimensional normed space* $(\mathcal{V}_i, \|\cdot\|_i) \cong (\mathbb{R}^{D_i}, \|\cdot\|_i)$.

Working with the whole population, we usually consider the Euclidean normed space

$$\mathcal{V} := \left(\prod_{i=1}^{N} \mathbb{R}^{D_i}, \| \cdot \| \right),$$

where:

$$\|\boldsymbol{x}\|^2 := \sum_i \|\boldsymbol{x}^{(i)}\|_i^2.$$

Suppose that the population action at time t is $\boldsymbol{x}_t \in \mathcal{X}$. The payoff/reward agent i received after $\boldsymbol{x}_t$ is given by:

$$\mathrm{u}_i(\boldsymbol{x}_t^{(i)}, \boldsymbol{x}_t^{(-i)}),$$

where $\mathrm{u}_i : \mathcal{X} \to \mathbb{R}$ is the utility function of the i. Throughout this paper, we mostly assume the following regularity condition for the utility functions:

Assumption 2.2: *For all $i \in [N]$ and $\boldsymbol{x}^{(-i)} \in \mathcal{X}_{-i}$, $\mathrm{u}_i((\cdot), \boldsymbol{x}^{(-i)})$ is concave and $\mathbf{v} := (\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(N)})$ is continuous where:*

$$\mathbf{v}^{(i)}(\boldsymbol{x}) := \nabla_{\boldsymbol{x}^{(i)}} \mathrm{u}^{(i)}(\boldsymbol{x}), \quad i \in [N].$$

A classical notion of equilibrium in games is the so-called Nash equilibrium:

Definition 2.3 (Nash Equilibrium): *Let be $\epsilon > 0$. $\overline{\boldsymbol{x}}_N \in \mathcal{X}$ is said to be an ϵ-Nash equilibrium of the game Γ if for every $i \in [N]$:*

$$\mathrm{u}_i(\overline{\boldsymbol{x}}_N) \geqslant \mathrm{u}_i(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{x}}_N^{(-i)}) - \epsilon, \quad \forall x^{(i)} \in \mathcal{X}_i \tag{2.4}$$

If above inequality is fulfilled with $\epsilon = 0$, then we say $\boldsymbol{x}_*$ is a Nash equilibrium.

So according to above definition, a Nash equilibrium describes the state in which no agent can increase his payoff by unilaterally changing his strategy.

Online Learning in a Competitive Environment

Due to the non-cooperativity assumption, a single rational agent in a repeated non-cooperative game faces the problem of sequential decision making in an unknown environment.

2.3. Game Theory and Variational Inequalities

This section aims to relate the concept of the Nash equilibrium of the NGCC to another alternative concept, which is more suitable for algorithmic analysis. For this sake, we introduce in the first part of this section the concept of variational inequality (VI), which is known to be a powerful and unifying method to study equilibrium problems both in infinite - and finite dimensions [55]. We will further see that the solutions to this inequality problem are Nash equilibriums. Furthermore, it is, in general, not comfortable working with a set that has no product structure, such as the feasible set of NGCC $\mathcal{Q}$. Thus in the second part of this section, we consider a Karush-Kuhn-Tucker-system-based technique to extend the corresponding VI to a VI on a set with a product structure (having possibly higher dimension) such that the solution structure remains preserved. The results stated in this chapter are known and can be found, for instance in [55].

Nash Equilibrium & Variational Inequality

The following concept is central to the analysis of first-order methods:

Definition 2.4 (Variational Inequality (VI)): *Let $\mathcal{Z}$ be a subset of a Euclidean normed space $\mathcal{V}$, and suppose that $\mathbf{g} : \mathcal{Z} \to \mathcal{V}^*$.*

- *A point $\overline{\boldsymbol{x}} \in \mathcal{Z}$ is a solution of the variational inequality VI($\mathcal{Z}, \mathbf{g}$), if :*

$$\langle \boldsymbol{x} - \overline{\boldsymbol{x}}, \mathbf{g}(\overline{\boldsymbol{x}}) \rangle \leqslant 0, \quad \forall \boldsymbol{x} \in \mathcal{Z}.$$

- *The set of solution of* $\mathrm{VI}(\mathcal{Z}, \mathbf{g})$ *is denoted by* $\mathrm{SOL}(\mathcal{Z}, \mathbf{g})$.

One helpful fact working with VI is that first-order methods tend to gain a drift toward the solution of it. For instance consider the iterates (4.2) and $\boldsymbol{x}_* \in \mathrm{SOL}(\mathcal{X}, \mathbf{v})$. If v is monotone (which is fulfilled by Assumption 2.2), then it follows that:

$$\langle \boldsymbol{X}_k - \boldsymbol{x}_*, \mathbf{v}(\boldsymbol{X}_k) \rangle \leqslant \langle \boldsymbol{X}_k - \boldsymbol{x}_*, \mathbf{v}(\boldsymbol{x}_*)) \rangle \leqslant 0, \tag{2.5}$$

where the last inequality follows from the definition of VI. Thus the first-order feedback $v(\boldsymbol{X}_k)$ for each time step $k+1$ forms an obtuse angle with the residual vector $X_k - \boldsymbol{x}_*$, and consequently, $v(\boldsymbol{X}_k)$ provides a direction toward $\boldsymbol{x}_*$.

There is no burden working with $\mathrm{SOL}(\mathcal{Q}, \mathbf{v})$ instead with GNE($\mathit{\Gamma}$) since, as asserted in the following proposition, the solutions of $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ is automatically a Nash equilibrium of $\mathit{\Gamma}$:

Proposition 2.5: *If Assumption 2.2 holds, then* $\mathrm{SOL}(\mathcal{Q}, \mathbf{v}) \subseteq$ *GNE($\mathit{\Gamma}$).*

Proof: The fact that a Nash equilibrium $\overline{\boldsymbol{x}}_N$ is in $\mathrm{SOL}(\mathcal{X}, \mathbf{v})$ follows from the first order condition of optimal point. That is since:

$$\overline{\boldsymbol{x}}_N^{(i)} = \arg\max_{x} \mathrm{u}(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{x}}_N^{(-i)}),$$

we have:

$$\langle \nabla_{\overline{\boldsymbol{x}}_N^{(i)}} \mathrm{u}(\overline{\boldsymbol{x}}_N), \overline{\boldsymbol{x}}_N^{(i)} - \boldsymbol{x}^{(i)} \rangle \geqslant 0, \quad \forall x^{(i)} \in \mathcal{X}_i.$$

If $\overline{\boldsymbol{x}}_N$ solves $\mathrm{VI}(\mathcal{X}, \mathbf{v})$, then:

$$\langle \mathbf{v}(\overline{\boldsymbol{x}}_N), x - \overline{\boldsymbol{x}}^N \rangle \geqslant 0, \quad \forall x \in \mathcal{X}.$$

Let be $i \in [N]$ arbitrary. Setting $x^{(j)} = x_N^{(j)}$, for all $j \in -i$, we have:

$$\langle \mathbf{v}^{(i)}(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{x}}_N^{-i}), \boldsymbol{x}^{(i)} - \overline{\boldsymbol{x}}_N^{(i)} \rangle \geqslant 0, \quad \forall x \in \mathcal{X}$$

Since $\mathbf{v}(\overline{\boldsymbol{x}}_N) = \nabla_{\overline{\boldsymbol{x}}_N^{(i)}} \mathrm{u}_i(\overline{\boldsymbol{x}}_N)$, u_i is concave in the i-th coordinate, and $\mathcal{X}_i$ is convex and non-empty, we have:

$$\mathrm{u}_i(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{x}}_N^{-i}) \leqslant \mathrm{u}_i(\overline{\boldsymbol{x}}_N) + \langle \mathbf{v}^{(i)}(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{x}}_N^{-i}), \boldsymbol{x}^{(i)} - \overline{\boldsymbol{x}}_N^{(i)} \rangle \leqslant \mathrm{u}_i(\overline{\boldsymbol{x}}_N),$$

and consequently:

$$\overline{\boldsymbol{x}}_N^{-i} = \operatorname*{arg\,min}_{\boldsymbol{x}^{(i)} \in \mathcal{X}_i} \mathbf{v}^{(i)}(\boldsymbol{x}^{(i)}, \overline{\boldsymbol{x}}_N^{-i}) \qquad \blacksquare$$

.

Remark 1: In the case where no coupling constraint is present, i.e., $\mathcal{C} = \mathcal{X}$, the converse of the above proposition holds. However, due to the coupling constraint, a Nash equilibrium has not to be a solution of the variational inequality.

This remark and Proposition 2.5 motivate us to highlight the generalized Nash equilibrium solving the corresponding variational inequality and call it as *variational Nash equilibrium.*

Another nice thing about VI is that under mild condition one can establish existence of its solution:

Proposition 2.6: *Let $\mathcal{Z}$ be a non-empty subset of a Euclidean normed space $\mathcal{V}$ and $\mathbf{g} : \mathcal{Z} \to \mathcal{V}^*$.*

1. *If that $\mathcal{Z}$ is compact and convex, and if $\mathbf{g}$ is continuous, then $\mathrm{SOL}(\mathcal{Z}, \mathbf{g}) \neq \varnothing$.*

2. *If $\mathrm{SOL}(\mathcal{Z}, \mathbf{g}) \neq \varnothing$ and $\mathbf{g}$ is strictly monotone on $\mathcal{Z}$, then $\mathrm{SOL}(\mathcal{Z}, \mathbf{g})$ is a singleton.*

Proof: Consider the mapping:

$$\mathbf{h} : \mathcal{X} \to \mathcal{X}, \ x \mapsto \mathbf{\Pi}_{\mathcal{X}}(\boldsymbol{x} - \mathbf{g}(\boldsymbol{x})),$$

where $\mathbf{\Pi}_{\mathcal{X}}$ denotes the projection onto $\mathcal{X}$. Clearly, $\mathbf{h}$ is continuous on the non-empty convex compact set $\mathcal{X}$, and $\mathbf{h}$ maps $\mathcal{X}$ into $\mathcal{X}$. Therefore by the Brouwer fixed point theorem, it follows that there exists a point $\overline{\boldsymbol{x}} \in \mathcal{X}$ s.t. $\overline{\boldsymbol{x}} = \psi(\overline{\boldsymbol{x}})$. This means that $\overline{\boldsymbol{x}} = \mathbf{\Pi}_{\mathcal{X}}(\overline{\boldsymbol{x}} - g(\overline{\boldsymbol{x}}))$ and consequently we have as desired:

$$-\langle g(\overline{\boldsymbol{x}}), y - \overline{\boldsymbol{x}} \rangle = \langle \overline{\boldsymbol{x}} - g(\overline{\boldsymbol{x}}) - \overline{\boldsymbol{x}}, y - \overline{\boldsymbol{x}} \rangle \leqslant 0, \quad \forall y \in \mathcal{X} \qquad \blacksquare$$

The proof of the first statement is based on the connection between VI and the fixed point of the Euclidean projection onto $\mathcal{Z}$. The latter can be described by means of Brouwer Fixed Point Theorem. The proof of second statement is closely related to the fact that if $\mathbf{g}$ is stricly monotone and $\overline{\boldsymbol{x}} \in \mathrm{SOL}(\mathcal{Z}, \mathbf{g})$, then:

$$\langle \boldsymbol{x} - \overline{\boldsymbol{x}}, \mathbf{g}(\boldsymbol{x}) \rangle \leqslant \langle \boldsymbol{x} - \overline{\boldsymbol{x}}, \mathbf{g}(\overline{\boldsymbol{x}} \rangle \leqslant 0, \quad \forall \boldsymbol{x} \in \mathcal{Z}, \boldsymbol{x} \neq \overline{\boldsymbol{x}} \tag{2.6}$$

with equality if and only if $\boldsymbol{x} = \overline{\boldsymbol{x}}$.

From here, it is immediate to infer the existence of a Nash equilibrium. Indeed if we set $\mathbf{g} = \mathbf{v}$ and $\mathcal{Z} = \mathcal{Q}$, then above proposition asserts the existence of a solution of $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ and correspondingly by Proposition 2.5 the existence of a Nash equilibrium of Γ:

Corollary 2.7: *Suppose that the Assumption 2.2 holds. Then the NGCC Γ has a Nash equilibrium.*

2.4. Elements of Martingale Theory

In this work we assume that a probability space $(\Omega, \Sigma, \mathbb{P})$, and $\mathbb{F} := (\mathcal{F}_n)_{n \in \mathbb{N}_0}$ a filtration therein i.e. $\mathbb{F}$ is a monotonicaly increasing sequence in Σ, are given. For ease of notations, we denote for each $n \in \mathbb{N}_0$ the conditional expectation $\mathbb{E}[\cdot|\mathcal{F}_n]$ given $\mathcal{F}_n$ simply by $\mathbb{E}_n[\cdot]$. Let $(\boldsymbol{M}_n)_n$ be a sequence of random variables taking values on a normed Euclidean space $(\mathbb{R}^D, \|\cdot\|)$. We say:

- $(\boldsymbol{M}_n)_n$ is adapted if $\boldsymbol{M}_n$ is $\mathcal{F}_n$-measureable
- $(\boldsymbol{M}_n)_n$ is predictable if $\boldsymbol{M}_n$ is $\mathcal{F}_{n-1}$-measureable.

Definition 2.5 ((super-,sub-)martingale): *We say $(\boldsymbol{M}_n)_n$ is a (resp. super-,sub-)martingale if:*

- *$(\boldsymbol{M}_n)_n$ is adapted,*
- *$(\boldsymbol{M}_n)_n$ is square-integrable, i.e. $\mathbb{E}[\|\boldsymbol{M}_n\|^2] < \infty$, $\forall n \in \mathbb{N}$,*
- *It holds $\mathbb{E}_n[\boldsymbol{M}_{n+1}] = \boldsymbol{M}_n$ (resp. $\leqslant$, $\geqslant$, instead of $=$) for all n.*

Given a martingale $(\boldsymbol{M}_n)_{n \in \mathbb{N}}$. One usually normalizes it by considering the martingale $(\boldsymbol{M}_n - \mathbb{E}_{n-1}[\boldsymbol{M}_n])_{n \in \mathbb{N}}$. This gives rise to the following definition:

Definition 2.6 (Martingale difference sequence): *Let $(\boldsymbol{M}_n)_{n \in \mathbb{N}}$ be a martingale. We say $(\boldsymbol{M}_n)_{n \in \mathbb{N}}$ is a martingale difference sequence if for all $n \in \mathbb{N}$:*

$$\mathbb{E}[\boldsymbol{M}_n|\mathcal{F}_{n-1}] = 0 \quad a.s.$$

One of the central result in the theory of martingales is the following result due to Doob:

Theorem 2.8 (Doob's Martingale Convergence Theorem): *Let $(X_n)_{n\in\mathbb{N}}$ be a real-valued super-martingale. If:*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[[X_n]_-\right] < \infty,$$

Then $(X_n)_{n\in\mathbb{N}}$ converges a.s. to a RV with a finite expectation.

A consequence of above theorem which will be used in this work is the following:

Theorem 2.9 (Theorem 2.18. in [56]): *Let:*

$$S_n = \sum_{k=1}^{n} X_k, \quad n \in \mathbb{N},$$

be a martingale, and $(U_n)_{n\in\mathbb{N}}$ be a non-negative predictable process.

1. *If :*

$$p \in [1, 2],$$

then:

$$\sum_{k=1}^{\infty} \frac{X_k}{U_k} \quad \text{converges a.s. on the set} \quad \left\{ \sum_{k=1}^{\infty} \frac{\mathbb{E}_k[|X_k|^p]}{U_k^p} < \infty \right\}, \tag{2.7}$$

and:

$$\lim_{n\to\infty} \frac{S_n}{U_n} = 0 \quad \text{a.s. on the set} \quad \left\{ \lim_{n\to\infty} U_n = \infty, \sum_{k=1}^{\infty} \frac{\mathbb{E}_k[|X_k|^p]}{U_k^p} < \infty \right\}.$$

2. *If:*

$$p > 2,$$

then both of the convergence statements in (2.7) and (1) hold on the set:

$$\left\{ \sum_{k=1}^{\infty} \frac{\mathbb{E}_k[|X_k|^p]}{U_k^{1+\frac{p}{2}}} < \infty \right\},$$

3. On the Convergence of Online Mirror Descent for Aggregative Games with Approximated Aggregates

Abstract: In this chapter, we consider a class of continuous games called aggregative games, in which the payoff of each player can be expressed to depend on his own action and the aggregate of the population's action. This class of games appears in a vast number of engineering fields, among others signal processing and communications. We study the case where every agent possesses no knowledge about the instantaneous change of the aggregate, and where he has in addition merely an estimate rather than the actual value of the aggregate. We give a mild sufficient condition on the agent's estimate of the aggregate depending on the step size, such that the corresponding mirror ascent based algorithm, suited with the framework of no-regret online learning, converges to the Wardrop equilibrium.

3.1. Introduction

Aggregative Games and Engineering Applications Competitive selfish agents appear as a model in a vast number of applications such as signal processing and communications (see e.g. [57]), smart grid [58–66], competitive markets [67], and congestion control for networks [68]. The famous concept of non-cooperative continuous game theory enables one to analyze such model. One class of games which we consider in this work is the class of aggregative games [69–71]. The feature of this is that the behaviour of every agent is influenced by both, his own strategy and a quantity which depends on the aggregate action of the entire population. This is in particular reflected in the fact that the payoff of every playing agent in such games can be written in the way such that it depends on its own action and a function, called aggregate function, of the aggregate of the action chosen by all players.

Aggregative games are capable of adressing large population problems where the behavior of each agent is not only affected by specific one-to-one effects. This sort of game has applications not only in economics, e.g. Cournot's market model and Bertrand's market model (see e.g. [70]), but also in a large number of engineering applications such as network traffic [68,72], wireless systems [73], electricity [74], and commodity markets [75].

Limits of Agent's Macroscopic View In some applications, each agent has no knowledge about how the aggregate instantaneously changes if he deviates from his strategy. This is e.g. in the case where a large number of agents are present and therefore no central instance is available, which collects and broadcasts the actual state of the whole population. Moreover, factors such as the presence of malicious attackers, hardware dependent disturbances, and agents' lack of global view leads to the phenomenon that the agents have only an estimate of the actual value of the aggregate. So in applications, the only resource each agent might have in order to maximize its payoff is by observing the instantaneous change of its utility function for a fixed estimated aggregate value.

Problem Statement and Our Contribution In this chapter, we are interested in the population's behavior of online learning non-cooperative agents using mirror ascent based algorithm in order to optimize his own yield. We aim to answer the following question:

To what extent does the agent's lack of macroscopic view influence the population's behaviour?

In order to attain the answer, we propose a novel model of agent's dynamic described by an online-mirror-ascent-based algorithm. In contrast to the commonly applied, the proposed algorithm uses in each step a partial gradient on a fixed estimated aggregate rather than the full gradient of the player's utility function. We present a mild sufficient condition on the aggregate estimates such that the population's dynamic steers into an equilibrium. In contrast to the usual notion of equilibrium in game theory, the Nash equilibrium, we work with the concept of Wardrop equilibrium since first-order evolution algorithms converging to the Nash equilibrium require full gradient information. In case that the utility function possesses some regularity condition, Wardrop equilibrium can be seen as an almost Nash equilibrium. The benefit of the result given in this work is that it provides quantitative guidance to design/choose a signal processing method for the agents such that the population converges to the game's equilibrium.

Relation to Prior Work

A deterministic analysis of online mirror ascent algorithm generally for games with continuous action set was given in [76]. In contrast to this work we consider more specifically aggregative games and assume that the agents have no complete information on the gradient of their utility functions. Nevertheless, the main focus of [76] is the case where the full gradient is disturbed by martingale noise.

Closely related to our work is [77]. Similar to our setting, they assume that the agent is uninformed about the full gradient and the actual value of the aggregate. However, in contrast to our work, the considered action space is provided with Euclidean structure so that the proposed algorithm is basically a projected gradient ascent. Nevertheless, the aim of their work is more specific, viz. to design a distributed algorithm for estimation of the aggregate which leads the projected gradient ascent to the Wardrop equilibrium. Our result allows to extend their method to the non-Euclidean case by means of the mirror ascent based algorithm.

The relation between Nash and Wardrop equilibrium has been extensively studied in economics literature (see e.g. [78, 79]) and in engineering literature (see e.g. [80–82]). Recent works on that aspect is given in [83, 84]. They differ from previous research by additionally considering coupling constraints between the agents. However the projected gradient method proposed to achieve the equilibrium relies on the existence of a centralized unit.

3.2. Model Description and Basic Notions

As the basic setting, we consider a non-cooperative game $\mathit{\Gamma} = ([N], \mathcal{X}, \mathrm{u})$ with $\mathcal{X}$ and u given as in Section 2.2. $\mathrm{u} = (\mathrm{u}_1, \dots, \mathrm{u}_N)$ played by a finite set of players $[N] = \{1, \dots, N\}$.

3.2.1. Aggregative Games

In this chapter, we are interested in the following specific class of games:

Definition 3.1 (Aggregative Games): *The game $\mathit{\Gamma}$ is said to be an aggregative game if there exists a subset $\mathcal{Y} \neq \varnothing$ of a finite dimensional normed space $(\tilde{\mathcal{V}}, \|\cdot\|_{\tilde{\mathcal{V}}})$, functions $\sigma : \mathcal{X} \to \mathcal{Y}$ and $\mathrm{g}_i : \mathcal{X}_i \times \mathcal{Y} \to \mathbb{R}$ s.t.:*

$$\mathrm{u}_i(\boldsymbol{x}) = \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x})),$$

In order to illustrate the concept of aggregative game, we give in the following some examples:

Example 11 (Mixed Extension of finite games): Let $\tilde{\mathit{\Gamma}} = ([N], \mathcal{A}, \tilde{\mathrm{u}})$ be a finite game. We consider the case that each player i can choose independent from other players a mixed strategy $\boldsymbol{x}^{(i)}$ from $\mathcal{A}_i$, i.e. $\boldsymbol{x}^{(i)}$ is an element of the probability simplex $\mathit{\Delta}(\mathcal{A}_i)$ over $\mathcal{A}_i$. The expected payoff is given by:

$$\mathrm{u}_i(\boldsymbol{x}) = \sum_{\alpha_1 \in \mathcal{A}_1} \cdots \sum_{\alpha_N \in \mathcal{A}_N} \tilde{\mathrm{u}}_i(\alpha_1, \dots, \alpha_N) \boldsymbol{x}^{(1)}_{\alpha_1} \cdots \boldsymbol{x}^{N}_{\alpha_N}.$$

In case that $\tilde{\mathrm{u}}_i$ takes form:

$$\tilde{\mathrm{u}}_i(\alpha_1, \ldots, \alpha_N) = \tilde{\mathrm{u}}_i^i(\alpha_i) + \tilde{\mathrm{u}}_i^{-i}(\alpha_{-i}),$$

for a certain functions $\tilde{\mathrm{u}}_i^i$ and $\tilde{\mathrm{u}}_i^{-i}$. It holds:

$$\tilde{\mathrm{u}}_i(\boldsymbol{x}) = \sum_{\alpha_i \in \mathcal{A}_i} \tilde{\mathrm{u}}_i^i(\alpha_i) \boldsymbol{x}_{\alpha_i}^{(i)} + \sum_{\alpha_{-i} \in \mathcal{A}_i} \tilde{\mathrm{u}}_i^{-i}(\alpha_{-i}) \prod_{j \in -i} \boldsymbol{x}_{\alpha_j}^{(j)}.$$

The game $\mathit{\Gamma} = ([N], (\mathit{\Delta}(\mathcal{A}_i))_{i\in[N]}, (c_i)_{i\in[N]})$ is an aggregative game with:

$$\mathrm{g}_i(\boldsymbol{x}^{(i)}, \boldsymbol{y}) = \sum_{\alpha_i \in \mathcal{A}_i} \tilde{\mathrm{u}}_i^i(\alpha_i) \boldsymbol{x}_{\alpha_i}^{(i)} + y,$$

and the aggregator:

$$\sigma(\boldsymbol{x}) = \sum_{\alpha_{-i} \in \mathcal{A}_i} \tilde{\mathrm{u}}_i^{-i}(\alpha_{-i}) \prod_{j \in -i} \boldsymbol{x}_{\alpha_j}^{(j)}$$

where for every $i \in [N]$, the utility function of the i-th agent can be written in the form:

$$\mathrm{u}_i(\boldsymbol{x}) = \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x})),$$

with $\sigma : \mathcal{X} \to \mathcal{Y}$ and $\mathrm{g}_i : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, where $\mathcal{Y} \neq \varnothing$ is a subset of a finite dimensional normed space $(\tilde{\mathcal{V}}, \|\cdot\|_{\tilde{\mathcal{V}}})$.

For sake of simplicity, we assume throughout this chapter the following:

Assumption 3.1: *g_i and σ are continuously differentiable.*

The instantaneous change of the payoff of each player is determined by the gradient operator:

$$\mathbf{G}^{\text{Nash}} = (\mathbf{G}_1^{\text{Nash}}, \ldots, \mathbf{G}_N^{\text{Nash}}) : \mathcal{X} \to \prod_{i=1}^{N} (\mathcal{V}_i^*, \|\cdot\|_{i,*}), \quad \boldsymbol{x} \mapsto \nabla_{\boldsymbol{x}^{(i)}} \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(-i)}))$$

In order to compute $\mathbf{G}^{\text{Nash}}$, each agent has to be able to anticipate the (instantaneous) change of the aggregative function. So for our purpose, it is beneficial to consider the operator:

$$\mathbf{G}^{\text{Ward}} = (\mathbf{G}_1^{\text{Ward}}, \ldots, \mathbf{G}_N^{\text{Ward}}) : \mathcal{X} \to \prod_{i=1}^{N} \mathcal{X}_i^*,$$

where:

$$\begin{aligned} &\mathbf{G}_i^{\text{Ward}}(\boldsymbol{x}) = \tilde{\mathbf{G}}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x})), \\ &\text{and} \quad \tilde{\mathbf{G}}_i : \mathcal{X}_i \times \mathcal{Y} \to \mathcal{X}_i^*, \quad (\boldsymbol{x}^{(i)}, \boldsymbol{y}) \mapsto \nabla_{\boldsymbol{x}^{(i)}} \mathrm{g}_i(\boldsymbol{x}^{(i)}, \boldsymbol{y}). \end{aligned} \tag{3.1}$$

3.2.2. Dynamic Model

In this chapter we are interested in gradient-based evolution. However, the gradient is in general an element of the dual space. So, in order to implement their actions based on this information, an agent needs a mapping to project the iterate back to his strategy set. A canonical way to do is by means of the mirror map given in the Definition 2.1. Throughout this chapter, we assume the following:

Assumption 3.2: *that each agent $i \in [N]$ possess a K_i-strongly convex regularizer ψ_i which induces the mirror map $\mathbf{\Phi}_i$.*

In order to emphasize the action of the whole population, we sometimes use the operator:

$$\mathbf{\Phi} : \mathcal{V}^* \to \mathcal{X}, \quad \boldsymbol{y} \mapsto (\mathbf{\Phi}_1(\boldsymbol{y}^{(1)}), \ldots, \mathbf{\Phi}_N(\boldsymbol{y}^{(N)})).$$

The iterate of the agents which we consider is specifically given in Algorithm 2. It is a

Algorithm 2 Online Mirror Ascent with Estimated Aggregation

Require: Step size sequence $(\gamma_n)_n$, initial dual action $\boldsymbol{Y}_0^{(i)} \in \mathcal{V}_i^*$
  for $n = 0, 1, 2, \ldots$ **do**
    for every player $i \in [N]$ **do**
      Play $\boldsymbol{X}^{(i)} \leftarrow \mathbf{\Phi}_i(\boldsymbol{Y}_n^{(i)})$
      Make an estimate $\hat{\sigma}_n^{(i)}$ of the actual aggregate $\sigma(\boldsymbol{X}_n)$
      Observe $\tilde{\mathbf{G}}(\boldsymbol{X}_n^{(i)}, \hat{\sigma}_n^{(i)})$
      Update $\boldsymbol{Y}_{n+1}^{(i)} \leftarrow \boldsymbol{Y}_n^{(i)} + \gamma_n \tilde{\mathbf{G}}(\boldsymbol{X}_n, \hat{\sigma}_n^{(i)})$
    end for
  end for

modification of the online mirror ascent algorithm (see [76]) fulfilling the no-regret policy, in which the iterate of agent $i \in [N]$ at the time $n \in \mathbb{N}$ takes the form:

$$\boldsymbol{X}_{n+1}^{(i)} = \mathbf{\Phi}_i(\boldsymbol{Y}_{n+1}^{(i)}),\ \boldsymbol{Y}_{n+1} = \boldsymbol{Y}_n + \gamma_n \mathbf{G}_i^{\mathrm{Nash}}(\boldsymbol{X}_n). \tag{3.2}$$

The difference between above algorithm and our algorithm lies in the gradient step: While the gradient step in (3.2) involves the full gradient $\nabla_{\boldsymbol{x}^{(i)}} \mathrm{g}_i(\boldsymbol{x}^i, \sigma(\boldsymbol{x}))$, the gradient step in algorithm 2 involves only the partial gradient $\nabla_{\boldsymbol{x}^{(i)}} \mathrm{g}_i(\boldsymbol{x}^i, \boldsymbol{y})|_{\boldsymbol{y}=\hat{\sigma}_n^{(i)}}$ evaluated at the estimate aggregate.

Remark 2: One possible method to estimate the actual aggregate, in case that it is a sum, is given in [77]. There, each agent distributedly estimates the population's aggregate by collecting the local neighbors' state.

3.3. Variational Description of Wardrop Equilibrium

3.3.1. Wardrop Equilibrium

A classical notion of equilibrium in games is the Nash equilibrium, which describes the state in which no agent can increase his payoff by unilaterally changing his strategy (see Definition 2.3). In the context of aggregative games, the inequality (2.4) specifies as:

$$\mathrm{g}_i(\boldsymbol{x}_*^{(i)}, \sigma(\boldsymbol{x}_*)) \geqslant \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_*^{(-i)})) - \epsilon.$$

The drawback of the concept of Nash equilibrium is that it requires that each agent has fully knowledge of his contribution to the aggregate function, which is of course not always the case. So, a reasonable replacement of this concept is the concept of Wardrop equilibrium:

Definition 3.2 (Wardrop Equilibrium): *$\boldsymbol{x}_* \in \mathcal{X}$ is a Wardrop equilibrium of the game Γ if for every $i \in [N]$:*

$$\mathrm{g}_i(\boldsymbol{x}_*^{(i)}, \sigma(\boldsymbol{x}_*)) \geqslant \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_*)), \quad \forall \boldsymbol{x}^{(i)} \in \mathcal{X}_i. \tag{3.3}$$

3.3.2. Existence Theorems for Wardrop Equilibria

For analysis purposes it is advantageous to relate the concept of Wardrop equilibrium to the concept of variational inequality (see Definition 2.4). For that sake, we assume the following:

Assumption 3.3: *For all $i \in [N]$:*

1. *$\mathrm{u}_i((\cdot), \boldsymbol{x}^{(-i)})$ is concave for all $\boldsymbol{x}^{(-i)} \in \mathcal{X}_{-i}$*

2. *$\mathrm{g}_i((\cdot), \boldsymbol{y})$ is concave for all $y \in \mathcal{Y}$*

The corresponding relation between those concepts yields from the first order optimality condition for convex optimization.

Proposition 3.1: *If for every $i \in [N]$, $\mathrm{g}_i((\cdot), \boldsymbol{y})$ is concave for all $y \in \mathcal{Y}$. Then the set of Wardrop equilibrium coincides with* $\mathrm{SOL}(\mathcal{X}, \mathbf{G}^{Ward})$

Proof: Suppose that $\boldsymbol{x}_W$ is a Wardrop equilibrium. By definition, we have:

$$\boldsymbol{x}_W^{(i)} = \underset{\boldsymbol{x}^{(i)} \in \mathcal{X}_i}{\arg\max}\, h_i(\boldsymbol{x}^{(i)}), \tag{3.4}$$

where:

$$h_i : \mathcal{X}_i \to \mathbb{R}, \quad \boldsymbol{x}^{(i)} \mapsto \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)).$$

Furthermore, we have:

$$\nabla_{\boldsymbol{x}^{(i)}} h_i(\boldsymbol{x}^{(i)}) = \nabla_{\boldsymbol{x}^{(i)}} \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)) = \tilde{\mathbf{G}}(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W))$$

So, the fact that $\boldsymbol{x}_W \in \mathcal{X}$ is a Wardrop equilibrium implies $\boldsymbol{x}_W$ is in $\mathrm{SOL}(\mathcal{X}, \mathbf{G}^{\mathrm{Nash}})$ follows from the first order condition of optimal point. That is (3.4) yields:

$$\langle \mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{x}_W^{(i)}), \boldsymbol{x}^{(i)} - \boldsymbol{x}_W^{(i)} \rangle = \langle \tilde{\mathbf{G}}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)), \boldsymbol{x}^{(i)} - \boldsymbol{x}_W^{(i)} \rangle \leqslant 0, \quad \forall \boldsymbol{x}^{(i)} \in \mathcal{X}_i. \tag{3.5}$$

Summing above inequality over all $i \in [N]$, we have that $\boldsymbol{x}_W \in \mathrm{SOL}(\mathcal{X}, \mathbf{G}^{\mathrm{Ward}})$.

For the converse, suppose that $\boldsymbol{x}_W$ solves $\mathrm{VI}(\mathcal{X}, \mathbf{G}^{\mathrm{Ward}})$, then:

$$\langle \mathbf{G}^{\mathrm{Ward}}(\boldsymbol{x}_W), \boldsymbol{x} - \boldsymbol{x}_W \rangle \leqslant 0, \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Let be $i \in [N]$ arbitrary. Setting $\boldsymbol{x}^{(j)} = \boldsymbol{x}_W^{(j)}$, for all $j \in -i$, we have:

$$\langle \mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{x}_W), \boldsymbol{x}^{(i)} - \overline{\boldsymbol{x}}_N^{(i)} \rangle \leqslant 0, \quad \forall \boldsymbol{x} \in \mathcal{X}. \tag{3.6}$$

Now, since $\mathrm{g}_i((\cdot), \sigma(\boldsymbol{x}_W))$ is concave in the i-th coordinate, and $\mathcal{X}_i$ is convex and non-empty, it follows that for any $\boldsymbol{y} \in \mathcal{Y}$:

$$\mathrm{g}_i(\boldsymbol{x}^{(i)}, \boldsymbol{y}) \leqslant \mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \boldsymbol{y}) + \langle \underbrace{\nabla_{\boldsymbol{x}_W^{(i)}} \mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \boldsymbol{y})}_{= \tilde{\mathbf{G}}_i(\boldsymbol{x}_W^{(i)}, \boldsymbol{y})}, \boldsymbol{x}^{(i)} - \boldsymbol{x}_W^{(i)} \rangle.$$

Setting $y = \sigma(\boldsymbol{x}_W)$ in above inequality, it yields:

$$\mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)) \leqslant \mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)) + \langle \mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{x}_W), \boldsymbol{x}^{(i)} - \boldsymbol{x}_W^{(i)} \rangle \leqslant \mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)), \tag{3.7}$$

where the inequality follows from (3.6). Finally, (3.7) asserts that for any $i \in [N]$:

$$\boldsymbol{x}_W^{(i)} \in \arg\max_{\boldsymbol{x}^{(i)} \in \mathcal{X}_i} \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)),$$

and thus $\boldsymbol{x}_W$ is a Wardrop equilibrium. ∎

Specifically if the 1st (resp. 2nd) condition in Assumption 3.3 holds, then the set of Nash (resp. Wardrop) equilibrium coincides with $\mathrm{SOL}(\mathcal{X}, \mathbf{G}^{\mathrm{Nash}})$ (resp. $\mathrm{SOL}(\mathcal{X}, \mathbf{G}^{\mathrm{Ward}})$). Thus to study the set of equilibrium, it is in our case sufficient to study the solution set of the corresponding variational inequality.

Now, assume that the utility function satisfies the following additional property:

Assumption 3.4: *For any $i \in [N]$, the family of functions $(\tilde{\mathbf{G}}_i(\boldsymbol{x}_i, (\cdot)))_{\boldsymbol{x}_i \in \mathcal{X}_i}$ is equicon-*

tinuous, in the sense that:

$$\forall \epsilon > 0: \quad \exists \delta > 0: \quad \forall \boldsymbol{x}^{(i)} \in \mathcal{X}_i: \quad \|\tilde{\mathbf{G}}_i(\boldsymbol{x}_i, \boldsymbol{y}) - \tilde{\mathbf{G}}_i(\boldsymbol{x}_i, \tilde{\boldsymbol{y}})\|_{i,*} < \epsilon, \quad \textit{if } \|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|_{\mathcal{V}_i} < \delta.$$

We can derive the existence and uniqueness of the Wardrop equilibrium from Proposition 3.1 as follows:

Proposition 3.2: *Suppose that the 2nd condition in Assumption 3.3 and Assumption 3.4 holds. Then Γ has a Wardrop equilibrium.*

Proof: To show the first statement, we need to show that:

$$\mathbf{G}^{\mathrm{Ward}} \text{ is continuous.} \tag{3.8}$$

From this result, we have by Proposition 2.6 that $\mathrm{SOL}(\mathcal{X}, \mathbf{G}^{\mathrm{Ward}})$ is non-empty and therefore by Proposition 3.1 also the set of Wardrop equilibrium. Toward this end, let be $i \in [N]$ arbitrary. Moreover, we fix $\boldsymbol{x} \in \mathcal{X}$. For any $\boldsymbol{z} \in \mathcal{X}$:

$$\begin{aligned} &\|\mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{x}) - \mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{y})\|_{i,*} = \|\tilde{\mathbf{G}}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x})) - \tilde{\mathbf{G}}_i(\boldsymbol{z}^{(i)}, \sigma(\boldsymbol{z}))\|_{i,*} \\ &\leqslant \|\tilde{\mathbf{G}}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x})) - \tilde{\mathbf{G}}_i(\boldsymbol{z}^{(i)}, \sigma(\boldsymbol{x}))\|_{i,*} + \|\tilde{\mathbf{G}}(\boldsymbol{z}^{(i)}, \sigma(\boldsymbol{x})) - \tilde{\mathbf{G}}_i(\boldsymbol{z}^{(i)}, \sigma(\boldsymbol{z}))\|_{i,*}. \end{aligned} \tag{3.9}$$

Now, let be $\epsilon > 0$ fixed. Since $\tilde{\mathbf{G}}_i((\cdot), \boldsymbol{y})$ is continuous for any $y \in \mathcal{Y}$, we can find $\delta > 0$ such that:

$$\|\tilde{\mathbf{G}}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x})) - \tilde{\mathbf{G}}_i(\boldsymbol{z}^{(i)}, \sigma(\boldsymbol{x}))\|_{i,*} < \frac{\epsilon}{2}, \quad \text{if}$$

for any $z \in \mathcal{X}$ satisfying:

$$\|\boldsymbol{x}^i - \boldsymbol{z}^{(i)}\|_i < \delta.$$

By the equicontinuity property, we can find $\tilde{\delta}$ such that for any $\boldsymbol{z}^{(i)} \in \mathcal{X}_i$:

$$\|\tilde{\mathbf{G}}(\boldsymbol{z}^{(i)}, \sigma(\boldsymbol{x})) - \tilde{\mathbf{G}}_i(\boldsymbol{z}^{(i)}, \boldsymbol{y})\|_{i,*} < \frac{\epsilon}{2},$$

whenever:

$$\|\sigma(\boldsymbol{x}) - \boldsymbol{y}\|_{\tilde{V}} < \tilde{\delta}.$$

Now, since σ is continuous, it follows that there exists $\delta_0 > 0$ s.t.:

$$\|\sigma(\boldsymbol{x}) - \sigma(\boldsymbol{z})\|_{\tilde{V}} < \tilde{\delta},$$

whenever:

$$\|\boldsymbol{x} - \boldsymbol{z}\| < \delta_0.$$

Combining all previous observations with (3.9), we obtain that:

$$\|\mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{x}) - \mathbf{G}_i^{\mathrm{Ward}}(\boldsymbol{y})\|_{i,*} < \epsilon,$$

whenever:

$$\|\boldsymbol{x} - \boldsymbol{z}\| < \min\{\delta, \delta_0\}, \qquad \blacksquare$$

which shows (3.8).

At last we present the following relation between Nash equilibrium and Wardrop equilibrium, which is an extension of Theorem 1 in [84]:

Proposition 3.3: *Suppose that for all $\boldsymbol{x}^i \in \mathcal{X}_i$:*

- $\mathrm{g}_i(\boldsymbol{x}^i, (\cdot))$ *is L_i^{g}-Lipschitz continuous*

- *for all $\boldsymbol{x}^{(-i)} \in \mathcal{X}_{-i}$, $\sigma((\cdot), \boldsymbol{x}^{(-i)})$ is L_i^{σ}-Lipschitz continuous,*

then it follows that a Wardrop equilibrium is a ϵ-Nash equilibrium with:

$$\epsilon = L_i^{\mathrm{g}} L_i^{\sigma} |\mathcal{X}_i|.$$

Proof: Let $\boldsymbol{x}_W$ be a Wardrop equilibrium Γ_N. It holds:

$$\mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)) \leqslant \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)), \quad \boldsymbol{x}^{(i)} \in \mathcal{X}_i. \tag{3.10}$$

Consequently, it yields:

$$\begin{aligned}
&\mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_W^{-i}))\\
&= \mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)) + \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_W^{-i}))\\
&= \mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_W^{-i})) + \underbrace{\mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W))}_{\leqslant 0 \text{ By (3.10)}}\\
&\leqslant \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_W^{-i})).
\end{aligned}$$

Since $\mathrm{u}_i(\boldsymbol{x}, (\cdot))$ and $\sigma(\cdot, \boldsymbol{x}^{-i})$ are continuous functions on compact sets $\mathcal{Y}$ and $\mathcal{X}_i$, it is Lipschitz. Therefore:

$$\begin{aligned}
\mathrm{g}_i(\boldsymbol{x}_W^{(i)}, \sigma(\boldsymbol{x}_W)) - \mathrm{g}_i(\boldsymbol{x}^{(i)}, \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_W^{-i})) &\leqslant L_i \|\sigma(\boldsymbol{x}_W) - \sigma(\boldsymbol{x}^{(i)}, \boldsymbol{x}_W^{(-i)})\| \leqslant L_i L_i^{\sigma} \|\boldsymbol{x}^{(i)} - \boldsymbol{x}_W^{(i)}\|\\
&\leqslant L_i L_i^{\sigma} |\mathcal{X}_i| \qquad \blacksquare
\end{aligned}$$

Remark 3: In the particular case where the aggregate is the mean of the population's action, i.e.:

$$\sigma(\boldsymbol{x}) = \frac{\sum_{i=1}^{N} \tilde{\sigma}(\boldsymbol{x}^{(i)})}{N}, \quad \text{where } \tilde{\sigma} \text{ is Lipschitz continuous,}$$

it yields from above prop. that a Wardrop equilibrium is a ϵ_N-Nash equilibrium, where $(\epsilon_n)_n$ is a decreasing sequence of non-negative numbers converging to 0.

3.4. Convergence Analysis

In this section we investigate the convergence behaviour of Algorithm 1 to the solution of the variational inequality $\mathrm{VI}(\mathcal{X}, \mathbf{G}^{\mathrm{Ward}})$. From the previous section it follows that if the condition 2 in Assumption 3.3 is fulfilled, the latter coincides with the set of Wardrop equilibria. To keep the argumentation short we assume throughout:

Assumption 3.5: $\mathbf{G}^{Ward}$ *is strictly monotone.*

This yields that (see Proposition 2.6):

$$\text{the solution } \{\boldsymbol{x}_*\} := \mathrm{SOL}(\mathcal{X}, \mathbf{G}^{\mathrm{Ward}}) \text{ is unique.}$$

We first measure the distance between the evolution of each agent and the equilibrium naturally by means of the "total" Fenchel coupling:

$$\overline{\mathrm{F}} : \mathcal{X} \times \prod_{i=1}^{N} \mathcal{V}_i^*, \quad (\boldsymbol{x}, \boldsymbol{y}) \mapsto \sum_{i=1}^{N} \mathrm{F}_i(\boldsymbol{x}_i, \boldsymbol{y}_i),$$

where for any $i \in [N]$, F_i is the Fenchel coupling (see Definition 2.2) corresponding to the mirror map $\boldsymbol{\Phi}_i$.

By 1 in Proposition 2.4 it follows that convergence w.r.t. $\overline{\mathrm{F}}$ implies the convergence of the iterate w.r.t. the underlying norm $\|\cdot\|$. For the analysis, the following converse property is advantageous:

Assumption 3.6: *For any $p \in \mathcal{X}$ and any sequence $(\boldsymbol{Y}_n)_n$ in $\mathcal{V}^*$, it holds:*

$$\boldsymbol{\Phi}(\boldsymbol{Y}_n) \to \boldsymbol{p} \quad \Rightarrow \quad \overline{\mathrm{F}}(\boldsymbol{p}, \boldsymbol{Y}_n) \to 0.$$

In order to measure the variation of distances between subsequent iterates and the equilibrium $\boldsymbol{x}_*$, we apply the statement 1 in Proposition 2.4 and the definition of the algorithm and obtain:

$$\begin{aligned}\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_{n+1}) &\leqslant \overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n) + \langle \boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n, X_n - \boldsymbol{x}_* \rangle + \frac{1}{2K}\|\boldsymbol{Y}_{n+1} - \boldsymbol{Y}_n\|_*^2 \\ &\leqslant \overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n) + \gamma_n \omega_n + \frac{\gamma_n^2}{2K_i}\zeta_n,\end{aligned}$$

where:

$$\omega_n := \langle \boldsymbol{X}_n - \boldsymbol{x}_*, \tilde{\mathbf{G}}(\boldsymbol{X}_n, \hat{\sigma}_n) \rangle, \quad \zeta_n := \|\tilde{\mathbf{G}}(\boldsymbol{X}_n, \hat{\sigma}_n)\|_*^2.$$

Borrowing the term from the literature of online learning, we define the total regret in the n-th step w.r.t. $\boldsymbol{x}_*$ by:

$$\mathrm{Reg}_n = \sum_{k=0}^{n} \gamma_k \omega_k + \frac{1}{2K} \sum_{k=0}^{n} \gamma_k^2 \zeta_k, \quad \text{where } K := \min_{i \in [N]} K_i.$$

By the inequality:

$$\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_{n+1}) - \overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_0) \leqslant \mathrm{Reg}_n$$

which results from telescoping the previous inequality, the meaning of this term becomes obvious, i.e., to investigate the convergence behaviour of the algorithm, we need to analyze the total regret.

In order to continue, we connect the total regret with the variational inequality $\mathrm{VI}(\mathcal{X}.\mathbf{G}^{\mathrm{Ward}})$ by expanding:

$$\omega_n = \xi_n + \psi_n,$$

where:

$$\begin{aligned}\psi_k &:= \langle \boldsymbol{X}_k - \boldsymbol{x}_*, \tilde{\mathbf{G}}(\boldsymbol{X}_k, \hat{\sigma}_k) - \tilde{\mathbf{G}}(\boldsymbol{X}_k, \sigma(\boldsymbol{X}_k))\rangle,\\ \xi_k &:= \langle \boldsymbol{X}_k - \boldsymbol{x}_*, \mathbf{G}^{\mathrm{Ward}}(\boldsymbol{X}_k)\rangle,\end{aligned}$$

which leads to:

$$\mathrm{Reg}_n = \sum_{k=0}^{n} \gamma_k \xi_k + \sum_{k=0}^{n} \gamma_k \psi_k + \frac{1}{2K} \sum_{k=0}^{n} \gamma_k^2 \zeta_k.$$

The following Lemma shows that under usual conditions on the step size and certain quality of the estimation of the aggregates, the population's iterates recur to every neighborhood of the solution of the variational inequality. This is clearly a necessary condition for the convergence:

Lemma 3.4: *Suppose that for all $i \in [N]$ and every $\boldsymbol{x}_i \in \mathcal{X}_i$, $\tilde{\mathbf{G}}_i(\boldsymbol{x}^{(i)}, (\cdot))$ is Lipschitz continuous. Let $(\gamma_n)_n \subset \mathbb{R}^+$ be a non-increasing sequence satisfying:*

$$\sum_{k=0}^{\infty} \gamma_k = \infty \quad \sum_{k=0}^{\infty} \gamma_k^2 < \infty. \tag{3.11}$$

Moreover, suppose that for all $i \in [N]$:

$$\sum_{k=0}^{\infty} \gamma_k \|\hat{\sigma}_k^{(i)} - \sigma(\boldsymbol{X}_k)\|_{\tilde{V}} < \infty. \tag{3.12}$$

Then:

1. *$(\zeta_n)_n$ is uniformly bounded,*

2. *$\sum_{k=0}^{\infty} \gamma_k |\psi_k| < \infty$,*

3. *There exists a subsequence $(\boldsymbol{X}_{n_k})_k$ of $(\boldsymbol{X}_n)_n$ s.t. $\boldsymbol{X}_{n_k} \to \boldsymbol{x}_*$ as $k \to \infty$.*

Proof: We show the 1st statement.. Let be $i \in [N]$. Since $\tilde{\mathbf{G}}_i(\boldsymbol{x}^i, (\cdot))$ is Lipschitz in the second argument, we have by additionally applying the triangle inequality:

$$\begin{aligned}\|\tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \hat{\sigma}_n^{(i)})\|_* &\leqslant \|\tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \sigma(\boldsymbol{X}_n)) - \tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \hat{\sigma}_n^{(i)})\|_* + \|\tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \sigma(\boldsymbol{X}_n))\|_* \\ &\leqslant L_i \|\sigma(\boldsymbol{X}_n) - \hat{\sigma}_n^{(i)}\|_{\tilde{\mathcal{V}}} + \|\tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \sigma(\boldsymbol{X}_n))\|_{i,*},\end{aligned}$$

where $L_i > 0$ is a constant. Now, since $\tilde{\mathbf{G}}_i$ and σ are continuous, and $\mathcal{X}$ is compact, $(\tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \sigma(\boldsymbol{X}_n)))_n$ is uniformly bounded. So (3.12) and (3.11) implies:

$$(\tilde{\mathbf{G}}_i(\boldsymbol{X}_n^{(i)}, \hat{\sigma}_n^{(i)}))_n \text{ is uniformly bounded and therefore also } (\zeta_n)_n. \tag{3.13}$$

The 2nd statement follows since:

$$|\psi_k| \leqslant C \sum_{i=1}^{N} \|\sigma(\boldsymbol{X}_k) - \hat{\sigma}_{i,k}\|_{\tilde{\mathcal{V}}}$$

for a $C > 0$. For the final statement, notice that:

$$\mathrm{Reg}_n = \tau_n \left(\frac{\sum_{k=0}^n \gamma_k \xi_k}{\tau_n} + \frac{\sum_{k=0}^n \gamma_k \psi_k}{\tau_n} + \frac{1}{2K} \frac{\sum_{k=0}^n \gamma_k^2 \zeta_k}{\tau_n} \right), \tag{3.14}$$

where:

$$\tau_n := \sum_{k=0}^{n} \gamma_k.$$

Let be:

$$U \text{ an arbitrary neighborhood of } \boldsymbol{x}_*.$$

Suppose that w.l.o.g. : $\boldsymbol{X}_n \notin U$ for all $n \in \mathbb{N}$. Combining this with the fact that $\mathbf{G}^{\mathrm{Ward}}$ is strictly monotone (see Proposition 2.6), it follows that:

$$\langle \mathbf{G}^{\mathrm{Ward}}(\boldsymbol{X}_n), \boldsymbol{X}_n - \boldsymbol{x}_* \rangle \leqslant -c, \quad \forall n \in \mathbb{N},$$

for a $c > 0$. This yields:

$$\mathrm{Reg}_n = -\tau_n \left(c - \frac{\sum_{k=0}^n \gamma_k \psi_k}{\tau_n} - \frac{1}{2K} \frac{\sum_{k=0}^n \gamma_k^2 \zeta_k}{\tau_n} \right).$$

The 1st (resp. 2nd) statement in this lemma and (3.11) gives that the 3rd (resp. 2nd) summand in the previous equality converges to 0 as n goes to infinity. Finally, since $\sum_{k=0}^{\infty} \gamma_k = \infty$, we have:

$$\mathrm{Reg}_n \to -\infty.$$

This contradicts the fact that:

$$\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_0) \text{ is finite.}$$

■

The desired convergence result is given in the following:

Theorem 3.5: *Suppose that the conditions given in Lemma 3.4 holds. Then* $(\boldsymbol{X}_n)_n$ *converges to* $\boldsymbol{x}_*$.

Proof: First, assume that $(\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n))_n$ converges. Combining this with 3rd statement in Lemma 3.4 and Assumption 3.6 asserts, which give $\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_{n_k}) \to 0$ as $k \to \infty$, we have $(\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n))_n$ converges to zero and thus by Assumption 3.6 we obtain the desired statement. We show now that previous assumption is true. Notice that by 3rd stmt in Prop. 2.4, by 1st statement in Lemma 3.4, and by the fact that $\xi_k \leqslant 0$ for all k, it holds that there exists $\tilde{C} > 0$ s.t. for all n:

$$\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_{n+1}) \leqslant \overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n) + \phi_n, \tag{3.15}$$

where $\phi_n := \gamma_n \psi_n + \gamma_n^2 \tilde{C}$. Define $Z_n := \overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n) - \sum_{k=0}^{n-1} \phi_k$, for all n. It follows from (3.15) by straightforward computation that $(Z_n)_n$ is a monotonically decreasing sequence, i.e. $Z_{n+1} \geqslant Z_n$ for all n. Moreover, by (3.11) and (3.12) we have that $\phi_k \geqslant 0$ and $\sum_{k=0}^{\infty} \phi_k < \infty$ and thus $(Z_n)_n$ uniformly bounded below i.e. there exists $c > 0$ s.t. $Z_n \geqslant -c$ for all n. Both previous observations give that $(Z_n)_n$ converges and consequently since $\sum_{k=0}^{\infty} \phi_k$ converges, we have as desired $(\overline{\mathrm{F}}(\boldsymbol{x}_*, \boldsymbol{Y}_n))_n$ converges. ∎

3.5. Numerical Simulation

We consider a Cornout game model for spectrum access in cognitive radio network (see e.g. [85]). The variable x_i denotes the bandwidth rented by the secondary user (SU) i which lies between 0 and 1 MHz, i.e. :

$$\mathcal{X}_i = [0, 1] \text{ MHz}.$$

Denote the receiver signal-to-noise ratio (SNR) of the receiver by $\gamma > 0$ and the threshold of the bit error rate by $\mathrm{BER}_{\mathrm{tar}}$. The utility function of SU $i \in [N]$ is given by:

$$\mathrm{u}_i(x) = Cx^{(i)} - \frac{N(x^{(i)})^2}{2} - \tau y(\sigma(x))^{\tau-1} x^{(i)},$$

where:

$$\sigma(\boldsymbol{x}) = \sum_{i=1}^{N} x^{(i)}, \quad C = \log_2\left(1 + \frac{1.5}{\ln(0.2/\mathrm{BER}_{\mathrm{tar}})}\right),$$

and $\tau > 1$ and $y > 0$ are some constants. We assume that the SUs apply Algorithm 2 with the mirror map:

$$\boldsymbol{\Phi}(\boldsymbol{z}) = \frac{\exp(\boldsymbol{z})}{1 + \exp(\boldsymbol{z})}$$

induced by the regularizer:

$$h(x) = x\ln(x) + (1-x)\ln(1-x),$$

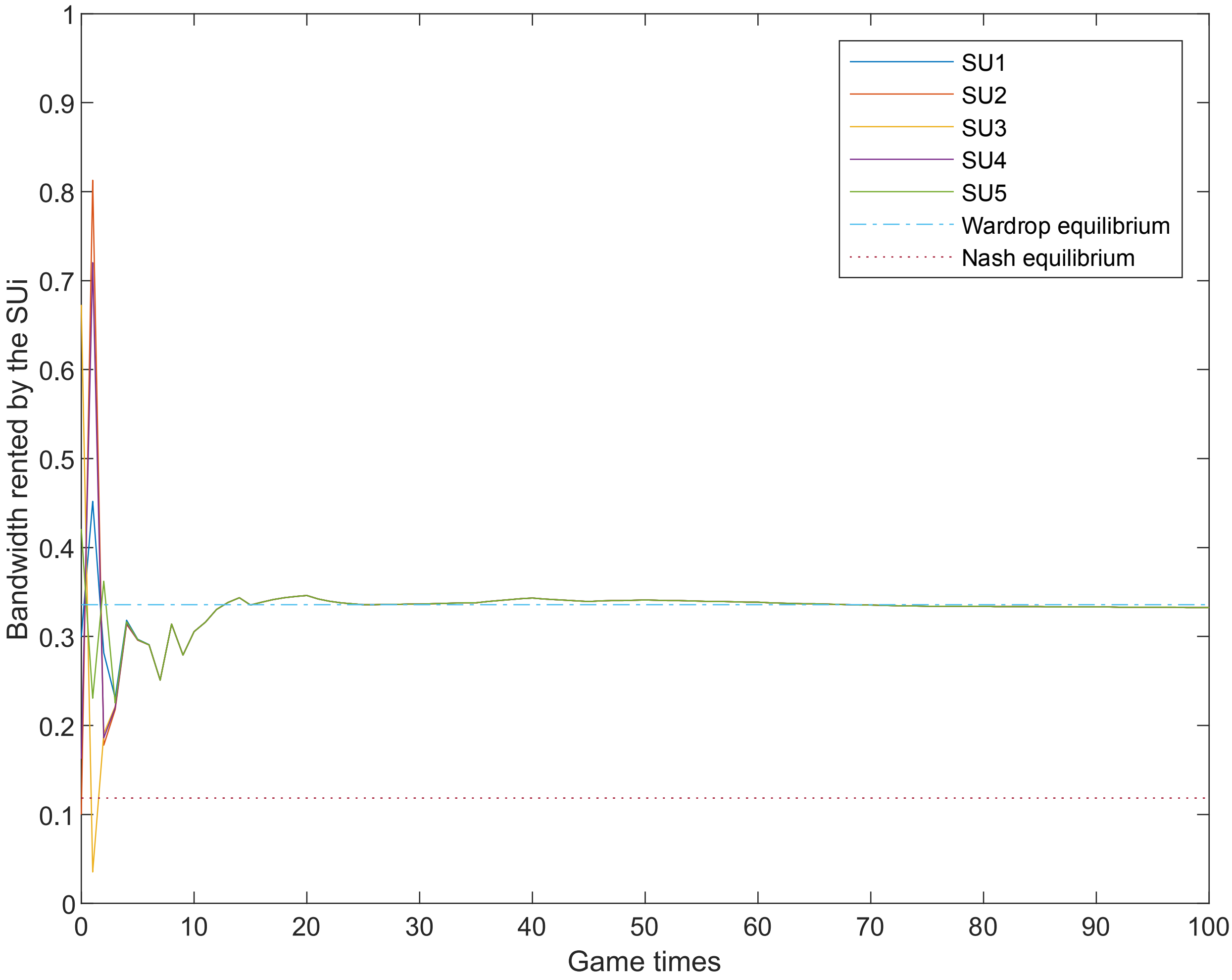


Figure 3.1.: Game Dynamic for 5 SUs

the step size sequence

$$\gamma_n = \frac{5}{n+1},$$

and the approximate aggregates:

$$\hat{\sigma}_n^{(i)} = \sigma(\boldsymbol{X}_n) + \frac{1}{n+1}.$$

We illustrate in Figure 3.1 the simulation for the case:

$$N = 5,\ \ y = 1, \tau = 2, \gamma = 15.4\text{dB}, \text{BER}_{\text{tar}} = 10^{-4},$$

and the initial rented bandwidth of each SUs chosen i.i.d. from $[0, 1]$. In Figure 3.1 it is apparent that the amount of bandwidth rented by the SUs converges to the Wardrop equilibrium which matches to our theoretical findings. We also simulate the cases where $N > 5$ and observe the same effect, although high N implies slower convergence of the algorithm to the Wardrop equilibrium.

4. Coordinated Online Learning for Multi-Agent Systems with Coupled Constraints and Perturbed Utility Observations

Abstract: Competitive non-cooperative online decision-making agents whose action impacts increase congestion of scarce resources constitute a model for widespread modern large-scale applications. To ensure sustainable resource behavior, we introduce a novel method whose aim is to steer the agents toward a stable population state, fulfilling the given coupled resource constraints. The proposed method comprises decentralized resource pricing via augmentation of the game's Lagrangian based on the resource loads. Assuming that the online learning agents have merely noisy first-order utility feedback, we show that for polynomial decaying agents' step size/learning rate, the population's dynamic almost surely converges to generalized Nash equilibrium. A particular consequence of the latter is the fulfillment of resource constraints in the asymptotic limit. Moreover, we investigate the finite-time quality of the proposed algorithm by giving a non-asymptotic time decaying bound for the expected amount of resource constraint violation.

4.1. Introduction

In a vast number of real-world applications (see also [86]) such as smart grid [58, 59, 61, 63–66], competitive markets [67], and network management [68], the view of the system participants as competitive selfish rational agents has become popular and led to fruitful discussions about system designs. Because of the emergence of several world-changing technologies, such as IoT, 5G, and the smart industry, and induced with them – the emergence of large-scale real-time systems, such view is becoming more and more important in the future. The reason is that the high-complexity, stringent latency -, and high-flexibility requirements in such systems, cause the communication and agreement between all their participants to be hard to establish, and therefore also the degree of cooperativeness between the agents to immensely decrease. This reason founds the importance of game theoretic modeling in the engineering.

Resource Constraints In widespread practical applications, the action of the agents causes the utilization of specific limited resources. For example: in network applications, the user's (agents) choice of data transfer paths (strategy) increases the congestion of specific links and routers (resources) with limited capacity; in electric mobility (see e.g. [63]), the vehicles' (agents) charge policy (strategy) increases the load of a grid, having limited electrical power (resource), at certain times; in fog networking (see e.g. [35]), the computation offloading choice (strategy) of a thing (agent) demands the computational power of certain fog nodes (resource). An important issue which has to be dealt by a system designer and - manager is the danger of resource overload due to agents' egoistic behavior, because the state of overutilization of resources can cause immense degradation of the overall system performance (see e.g. the problem of congestion and congestion collapse in networked system [87]) and negative environmental issues (e.g. caused by high CO2 emissions of electrical energy driven resources). Another example of events justifying the importance of the sustainability aspect in a system of egoistic optimizing agents is the flash crash in US financial markets due to fully automated computerized trading (see e.g. [88]).

Problem Description This work addresses the problem of how to control egoistic game-playing online learning agents, such that in the long-term, population's action converges to a stable state fulfilling the resource constraints. A challenge associated with this issue is to design a decentralized congestion control method that does not provide direct commands to each agent by a centralized instance and that it demands as little information about agents' characteristics as possible. The reason is that the methods contrary to the latter requirements would need, in case the number of agents is massive, exceptionally high computational power for the processing of the obtained information and the generation of the corresponding policies. Moreover, such methods would be inflexible for the possible exit of - and the entrance of new agents and therefore unsuitable for modern systems such as IoT.

Our Contributions Our main contribution is a novel method that solves the control problem as mentioned earlier in case the population of (online learning) selfish agents having disturbed first-order utility information. Its core is a resource pricing method aiming both, to give incentives (rather than direct commands/control) to all agents for acting sustainably, and to provide stability for the population's state. Our pricing method requires the current congestion state of the resources and not the specific characteristics of the agents. Moreover, it is done by the resources themselves rather than by a centralized instance.

Assuming for simplicity that the feedback noise is persistent.

- We give a sufficient condition on the agents' step size and the amount of augmentation such that the population iterate a.s. fulfills the resource constraints in the asymptotic limit. As a particular result it follows that if the agents' step size γ_n is of order $\Theta(n^{-\mu})$ where $\mu \in (1/2, 1]$, the population iterate converges almost surely to a (variational) Nash equilibrium of the corresponding game underlying coupled resource constraints and consequently almost surely fulfillment of the resource constraints in the asymptotic limit.

- We provide a non-asymptotic bound for the expected violation of the resource constraints. In particular, we show that if the agents' step size sequence is of order $\gamma_n = \Theta(n^{-1/2})$, the worst-case expected violation is of order $\mathcal{O}(\ln^{3/2}(n)/\sqrt{n})$.

- We are able to show that for a large class of decaying step size sequences of order $\gamma_n = \Theta(n^{-\mu})$ where $\mu \in (0, 1]$, the ergodic average of the population's iterate almost surely fulfills the resource constraints in the asymptotic limit.

- We provide a non-asymptotic bound for the expected distance of the iterations' ergodic average and the corresponding variational Nash equilibrium. In case that the noise is persistent, we show that for a fixed time horizon n and a certain fixed step size, bound of order $\mathcal{O}(n^{-1/2})$ is achievable. Moreover, we show that expectation bound of order $\mathcal{O}(G^2 \ln(n) n^{-1/2})$ is achievable if instead variable step-size of order $\mathcal{O}(n^{-1})$ is used.

- Under the additional assumption that the occurring noise has a light tail, we are even able to provide a high probability bound for the aforementioned distance. In particular, we show that a bound of order ϵ is achieved with probability $1 - \delta$ for the time horizon $n \geqslant \mathcal{O}(\ln(2/\delta)/\epsilon^2)$by a suitable choice of fixed step sizes.

The proposed method can be used by a system designer to develop agents control algorithms, based on intrinsically motivated reinforcement learning, aiming to generate a desirable collective behavior (in case that the latter coincides with the Nash equilibrium of the considered coupled constrained game).

Relation to Prior Works

Learning in Games Our work is related to the works investigating the dynamic of learning agents in a competitive setting: Several works generate long-term results concerning different agent's types ranging from no-learning (e.g., greedy agents – best-response dynamic) to learning agents, such as the fictitious play, where the agents keep track only of data of opponents play, the gradient play, where the agents learn from the first-order feedback of their utility. For a comprehensive review of the literature on those topics, we refer to [89]. Also, recent works analyze farsighted reinforcement learning agents (for a recent survey, see [90]). Here, we focus on online learning agents applying the canonical mirror descent algorithm [47, 48]. Therefore, the closest work to ours is [76] (along with several extensions such as [91,92]). In contrast to [76], our aim is not to predict the long-term outcome of a game with online learning agents, but rather to control competitive online learning agents respective to coupled resource constraints. Therefore in our case, the admissible set of population strategy profiles is not necessarily of product structure. For this reason, we have to modify the decentralized algorithm given in [76] and make use of coordinators to handle such inter-agent constraints. Moreover, in contrast to [76], where continuous interpolation is used to show the convergence of the given discrete-time algorithm, our technique relies directly on the martingale convergence theorem.

Generalized Nash equilibrium and Coupled Constraints As we consider a non-cooperative game with coupled constraints, we mention in the following some related works on the generalized Nash equilibrium. Usually of interest is the subclass of the variational Nash equilibrium, defined as the solution of the well-known concept of the variational inequality [55]. Several characterizations of the Nash equilibrium of games with coupled (resource) constraints have been made in the works [55, 93–97] leveraging from the duality theory and the theory of variational inequalities. Our work does not overlap with those mentioned work as our emphasis is not on analyzing the generalized Nash equilibrium. In general, a Nash equilibrium does not coincide with the population's welfare and, therefore, not efficient. There is extensive literature (e.g., [98–100]) on quantifying the loss of efficiency resulted from the population being in a Nash equilibrium. Relevant to our work is the very recent analysis of the efficiency loss in a non-cooperative game with coupled constraints [101]. [101] gives a hint that the variational Nash equilibrium might be efficient, but the generalized Nash equilibrium might be arbitrarily inefficient. So a method converging to a variational Nash equilibrium might not only support resource sustainability but also increases the population's welfare. In our numerical simulation, we also observe this effect.

Nash Equilibrium Findings Our method is also suitable for the Nash equilibrium finding. There is a large body of literature considering this problem. Reviewing them is beyond the scope of this work. Thus we concentrate on those that consider a similar setting as ours, i.e., the game with coupled constraints. Most of the existing works such as [84, 95, 102] follow this approach and proposes a primal-dual algorithm based on the fixed-point methods for finding the solution of a variational inequality (see, e.g., Chapter 12 in [55]), resulting in a Euclidean-projection based algorithm. In contrast to the method given in the mentioned work, our method uses the mirror map, which constitutes a generalization of the Euclidean projection. As already noticed in [48], the advantage of using a mirror map other than Euclidean projection that the algorithm performance of the former might have a weaker dependence on the underlying dimension of the decision space. Since we use the mirror map, we are not able to use the usual convergence proof via a fixed-point approach for the variational inequality. Furthermore, another essential difference between our algorithm and first-order algorithms for finding Nash equilibria of a game with coupled constraints is that they mostly use constant step size, do not consider the possibility of noise in the feedback information, and do not have a non-asymptotic guarantee of the violation of the resource constraints.

Also worth mentioning are works that consider the payoff-based approach (see, e.g., [103,104] and the references therein), where each agent can only observe its obtained payoffs. Such an approach is important for some applications, e.g., [105,106]. In contrast, our work assumes that the agents have each gradient observation of their utility. Nevertheless, investigating a method for gradient feedback constitutes a cornerstone for a payoff-based approach. So, we expect that from our work, one can generate a control-method via pricing for resource sustainability for online learning payoff-based agents.

Resource Congestion Control At last, we mention that the problem of alleviation of resource congestion can also be combined with other objectives such as the maximization of the population's welfare. Such a practice is common for instance in the field of network/internet congestion control [107], see also e.g., [108] and the control theoretic approach [109]. However, most of the algorithms for fulfilling such extended task requires either high degree of control of the agents, or specific information about the agents (e.g., their utilities). In contrasts our method only assumes that the agents are online learner and based of pricing instrument via resource congestion state. However, the cost we pay is that we cannot guarantee (theoretically) performance gains other than the alleviation of resource congestion. However, practically, we can see additional gain of population's wealth (see Section 4.8).

Chapter Organization

The structure of our work is as follows. In Section 4.2, we set up the multi-agent setting of our consideration by introducing the underlying game and the notion of coupled (resource) constraints. Assuming that the agents' learning model in the repeated game setting is the online mirror ascent, we propose in Section 4.3 a pricing algorithm aiming to lead the agents to a stable resource sustainable state. The remaining sections are devoted to the analysis of our proposed method:

- In Section 4.4, we show that our method can to ensure resource sustainability in the asymptotic region by showing the convergence of the population's iterate to a stable set satisfying coupled constraints. Since the proof of our result is quite technical, we part it into several Subsections, which one can skip at the first reading.

- In Section 4.5, we quantify to what extent our method can reduce the amount of coupled constraint violations by deriving a time-decaying bound for the constraint violation caused by the price-controlled online learning population.

- In Section 4.6, we close the gap between parameter choices provided in Sections 4.4 and 4.5, and derive the convergence the ergodic average of the MDAL iterates to the stable state of the interest.
- In Section 4.7 we quantify non-asymptotically the convergence behaviour of MDAL by providing a bound for the distance of the iterate to the aforementioned variational Nash equilibrium.
- The final section (Section 4.8) is devoted to practical simulations. Not only we provide numerical support of our theoretical findings, but also we show that our method might not be too conservative since it can also ensure the increase of the population's wealth. Furthermore, we compare our method with the states of the art and show that it may outperform them.

4.2. Model Description and Preliminaries

The main objective of this section is twofold. First, to formalize the setting of competitive agents by introducing the notion of a non-cooperative game. Second, to define the so-called mirror map, which provides a model of how a selfish agent realizes her decision from the first-order feedback.

4.2.1. Coupled Resource Constraints

For a certain number $R > 0$ of resources, we model the relation between agents' action and resource utilization by a function $\mathbf{g} : \mathbb{R}^D \to \mathbb{R}^R$. Throughout this paper, we assume that the function $\mathbf{g}$ is subject to the following conditions

Assumption 4.1: *For all $r \in [R]$, $\mathbf{g}_r : \mathcal{X} \to \mathbb{R}$ is convex and differentiable, and the Jacobian matrix $\nabla \mathbf{g}$ of $\mathbf{g}$ is continuous*

One may interpret the term $\mathbf{g}_r(\boldsymbol{x})$ as the overload/congestion state of the resource $r \in [R]$ caused by the population action $\boldsymbol{x}$. Since from operational - and sustainability point of view overload has to be kept low and even avoided, it is desired that the population strategy is contained in:

$$\mathcal{Q} := \mathcal{C} \cap \mathcal{X},$$

where:

$$\mathcal{C} := \{\mathbf{g}(\boldsymbol{x}) \leqslant 0\}$$

denotes the resource constraints. In order that this goal is feasible, we assume that $\mathcal{C}$ is non-empty. The following regularity condition on $\mathcal{Q}$ is useful for later purposes:

Assumption 4.2 (Slater's condition): *There exists a point $\boldsymbol{x}_*$ in $\operatorname{relint}(\mathcal{X})$ s.t.:*

$$\mathbf{g}(\boldsymbol{x}_*) < 0,$$

where relint($\mathcal{X}$) denotes the relative interior of $\mathcal{X}$.

The constraint $\boldsymbol{x} \in \mathcal{C}$ is also known as the *coupled constraint*. The reason is that the compliance depends on the strategy, not only of a single agent but also of the whole population. NG, which is also subject to coupled inequality constraints $\mathcal{C}$, is also called NG with coupled constraints (NGCC). The set of the feasible strategy of the player i given a joint action $\boldsymbol{x}^{(-i)}$ of other agents is denoted by $\mathcal{Q}^{(i)}(\boldsymbol{x}^{(-i)}) := \{\boldsymbol{x}^{(i)} \in \mathcal{X}_i : \mathbf{g}(\boldsymbol{x}) \leqslant 0\}$.

4.2.2. Generalized Nash Equilibrium (GNE)

One of the central concepts in game theory is the so-called Nash equilibrium, which denotes a feasible strategy profile at which no agent can improve his reward by unilaterally deviating from his strategy. For general games possibly underlying besides coupled constraints, we formally define this notion as follows (see e.g. [93]):

Definition 4.1 (Generalized Nash Equilibrium (GNE)): *Given a NGCC Γ and a resource constraint $\mathcal{C}$. $\overline{\boldsymbol{x}} \in \mathcal{Q}$ is said to be a (generalized) Nash equilibrium of Γ with $\mathcal{C}$ if for each $i \in [N]$:*

$$\mathrm{u}_i(\boldsymbol{x}_*^{(i)}, \boldsymbol{x}_*^{(-i)}) \geqslant \mathrm{u}_i(\boldsymbol{x}^{(i)}, \boldsymbol{x}_*^{(-i)}), \quad \forall \boldsymbol{x}^{(i)} \in \mathcal{Q}(\boldsymbol{x}_*^{(-i)}).$$

The set of all generalized Nash equilibrium of Γ is denoted by GNE(Γ)

4.3. Mirror Ascent with Augmented Lagrangian

Throughout this work, we assume that each agent $i \in [N]$ possesses a K_i-strongly convex regularizer ψ_i that induces the mirror map $\boldsymbol{\Phi}_i$ and the Fenchel coupling F_i. For the sake of simplicity, we assume $K_i = K$, for all $i \in [N]$.In order to emphasize the action of the whole population, we use the operator $\boldsymbol{\Phi} : \mathcal{V} \to \mathcal{X},\ \boldsymbol{y} \mapsto (\boldsymbol{\Phi}_1(\boldsymbol{y}^{(1)}), \ldots, \boldsymbol{\Phi}_N(\boldsymbol{y}^{(N)}))$ and the total Fenchel coupling $\overline{\mathrm{F}} : \mathcal{X} \times \mathcal{V}^* \to \mathbb{R}_{\geqslant 0},\ (\boldsymbol{x}, \boldsymbol{y}) \to \sum_i \mathrm{F}_i(\boldsymbol{x}_i, y_i)$. Proposition 2.4 implies that "convergence" of the sequence $(\boldsymbol{\Phi}(\boldsymbol{Y}_n))_n$ induces by $(\boldsymbol{Y}_n)_n$ in the sense that $\overline{\mathrm{F}}(\boldsymbol{p}, \boldsymbol{Y}_n) \to 0$ for an $\boldsymbol{p} \in \mathcal{X}$ implies convergence of $(\boldsymbol{\Phi}(\boldsymbol{Y}_n))_n$ w.r.t. the underlying norm. For later purposes, it is helpful to assume that the converse statement holds:

Assumption 4.3 (Reciprocity Condition): *For any $\boldsymbol{p} \in \mathcal{X}$ and any sequence $(\boldsymbol{Y}_n)_n$ in $\mathcal{V}^*$, it holds: $\boldsymbol{\Phi}(\boldsymbol{Y}_n) \to \boldsymbol{p} \ \Rightarrow\ \overline{\mathrm{F}}(\boldsymbol{p}, \boldsymbol{Y}_n) \to 0$.*

The reciprocity condition is standard in the literature of mirror descent (see [110]).

The foundation of the algorithm proposed in this work is given by the following iterate of the agent $i \in [N]$:

$$\boldsymbol{X}_{n+1}^{(i)} = \boldsymbol{\Phi}_i(\boldsymbol{Y}_{n+1}^{(i)}),\ \boldsymbol{Y}_{n+1}^{(i)} = \boldsymbol{Y}_n^{(i)} + \gamma_n \mathbf{v}^{(i)}(\boldsymbol{X}_n), \tag{4.2}$$

Algorithm 3 Mirror Ascent Augmented Resource Pricing (MAARP)

Require: Step size sequence (γ_n), augmentation functions (θ_n)

Require: Initial dual action $\boldsymbol{Y}_0^{(i)} \in \mathcal{V}_i^*$, - dual variable $\boldsymbol{\Lambda}_0 \in \mathbb{R}_{\geqslant 0}^R$

for $n = 0, 1, 2, \ldots$ **do**

Population play $\boldsymbol{X}_n = \boldsymbol{\Phi}(\boldsymbol{Y}_n)$

for every player $i \in [N]$ **do**

Observe the gradient utility feedback

$$\hat{\mathbf{v}}_n^{(i)} := \mathbf{v}^{(i)}(\boldsymbol{X}_n) + \boldsymbol{M}_{n+1}^{(i)}$$

Query the gradient load feedback:

$$\nabla_{\boldsymbol{X}_n^{(i)}} \mathbf{g}_r(\boldsymbol{X}_n), \quad r \in [R]$$

Update the score vector

$$\boldsymbol{Y}_{n+1}^{(i)} \leftarrow \boldsymbol{Y}_n^{(i)} + \gamma_n \left(\hat{\mathbf{v}}_n^{(i)} - \sum_{r=1}^{R} \boldsymbol{\Lambda}_n^r \nabla_{\boldsymbol{X}_n^{(i)}} \mathbf{g}_r(\boldsymbol{X}_n) \right)$$

end for

for every resource $r \in [R]$ **do**

Check its load $\mathbf{g}_r(\boldsymbol{X}_n)$

Update the price:

$$\boldsymbol{\Lambda}_{n+1}^r \leftarrow \boldsymbol{\Pi}_{\mathbb{R}_{\geqslant 0}} \left(\boldsymbol{\Lambda}_n^r + \gamma_n \left[\mathbf{g}_r(\boldsymbol{X}_n) - [\nabla_{\boldsymbol{\Lambda}_n} \theta_n(\boldsymbol{\Lambda}_n)]_r \right] \right) \tag{4.1}$$

Broadcast $\boldsymbol{\Lambda}_{n+1}$ to all players.

end for

end for

where $\boldsymbol{X}_n = (\boldsymbol{X}_n^{(i)})_i$. This algorithm is a canonical extension of the standard mirror ascent algorithm within the framework of online learning [38] in which the learner subsequently tries to optimize his apriori unknown time-variant regret/payoff-function using some problem-specific feedback which is in our case the first-order information of his utility function. For practical reasons, we assume that each agent i does not know the exact gradient $\mathbf{v}^{(i)}$. We can model this aspect, by modifying (4.2) as follows:

$$\boldsymbol{X}_{n+1}^{(i)} = \boldsymbol{\Phi}_i(\boldsymbol{Y}_{n+1}^{(i)}),\ \boldsymbol{Y}_{n+1}^{(i)} = \boldsymbol{Y}_n^{(i)} + \gamma_n \hat{\mathbf{v}}_n^{(i)}, \tag{4.3}$$

where:

$$\hat{\mathbf{v}}_n^{(i)} = \mathbf{v}^{(i)}(\boldsymbol{X}_n) + \boldsymbol{M}_{n+1}^{(i)},$$

and $(\boldsymbol{M}_n^{(i)})_{n\in\mathbb{N}}$ be a $\mathcal{V}_i^*$-valued $\mathbb{F}$-martingale difference sequence (see Definition 2.6.).

Remark 4 (Justification of Martingale Noise model): This noise model is quite general. For instance it clearly covers the case where $(\boldsymbol{M}_n^{(i)})_n$ is an i.i.d. zero mean square-integrable noise. Moreover it covers also noises with memory, such as e.g.:

$$\boldsymbol{M}_n^{(i)} = \epsilon_n^{(i)}\epsilon_{n-1}^{(i)},$$

where $(\epsilon_n^{(i)})_n$ i.i.d. mean zero RV and $\mathbb{F}$ is the corresponding filtration (containing the filtration) generated by the history of $(\epsilon_n^{(i)})$. Alternative way to define $\hat{\mathbf{v}}_n^{(i)}$ is by requiring:

$$\mathbb{E}[\|\hat{\mathbf{v}}_n^{(i)}\|_{i,*}^2] < \infty \quad \text{and} \quad \mathbb{E}[\hat{\mathbf{v}}_n^{(i)}|\mathcal{F}_n] = \mathbf{v}^{(i)}(\boldsymbol{X}_n).$$

In particular if $\mathbb{F}$ is the filtration generated by the history of the iterates. So latter requirement $\mathbb{E}[\hat{\mathbf{v}}_n^{(i)}|\mathcal{F}_n] = \mathbf{v}^{(i)}(\boldsymbol{X}_n)$ means that given the history of the iterates until time n, $\hat{\mathbf{v}}_n^{(i)}$ is an unbiased estimator of $\mathbf{v}^{(i)}(\boldsymbol{X}_n)$. The corresponding martingale difference sequence is given by:

$$\boldsymbol{M}_{n+1}^{(i)} = \mathbb{E}[\hat{\mathbf{v}}_n^{(i)}|\mathcal{F}_n] - \mathbf{v}^{(i)}(\boldsymbol{X}_n).$$

In order to handle the coupled resource constraints, we design a pricing mechanism based on the augmented Lagrangian method (see e.g. [111]), which is done by the resources themselves based on their congestion state. The corresponding method is given more in detail in Algorithm 3. The population's iterate given in Algorithm 3 can be shortly written as:

$$\begin{aligned}\boldsymbol{X}_{n+1} &= \boldsymbol{\Phi}(\boldsymbol{Y}_{n+1}),\ \boldsymbol{Y}_{n+1} = \boldsymbol{Y}_n + \gamma_n \left(\hat{\mathbf{v}}_n - [\nabla \mathbf{g}(\boldsymbol{X}_k)]^{\mathrm{T}} \boldsymbol{\Lambda}_k\right) \\ \boldsymbol{\Lambda}_{n+1} &= \boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}} \left(\boldsymbol{\Lambda}_n + \gamma_n \left[\mathbf{g}(\boldsymbol{X}_n) - \nabla_{\boldsymbol{\Lambda}_n}\theta_n(\boldsymbol{\Lambda}_n)\right]\right)\end{aligned} \tag{4.4}$$

In this work we mainly consider the augmentation functions of the form:

$$\theta_n(\boldsymbol{\lambda}) := \alpha_n \|\boldsymbol{\lambda}\|_2^2/2, \quad \text{where } \alpha_n > 0. \tag{4.5}$$

Remark 5 (Decentralization Aspect): Suppose that coupled constraint is affine, i.e. the function $\mathbf{g}$ is given as:

$$\mathbf{g}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x} - b,$$

where:

$$\mathbf{A} \in \mathbb{R}^{R \times \sum_{i=1}^{N} D_i} \quad \text{and} \quad b \in \mathbb{R}^R.$$

Moreover denote:

$$\mathbf{A} := [\mathbf{A}_1, \ldots, \mathbf{A}_N],$$

where for all $i \in [N]$:

$$\mathbf{A}_i \in \mathbb{R}^{R \times D_i}$$

is the matrix which is available only to the agent i. In this setting, the iterate of agent i is given specifically by

$$\boldsymbol{X}_{n+1}^{(i)} = \boldsymbol{\Phi}_i(\boldsymbol{Y}_n^{(i)} + \gamma_n[\hat{\mathbf{v}}^{(i)} - \mathbf{A}_i^{\mathrm{T}} \boldsymbol{\Lambda}_n]).$$

Therefore the iterate of each agent at all times $n+1$ is decentralized in the sense that it only requires local information available for each agent, such as first-order feedback $\hat{\mathbf{v}}_n^{(i)}$ and the constraint matrix $\mathbf{A}_i$, and the prices $(\boldsymbol{\Lambda}_n^r)_{r \in [R]}$, set by the resources $r \in [R]$.

Remark 6 (Differences between MAARP and Alg. 2 in [84]): Besides the fact that the feedback in MAARP is noisy, a difference between MAARP and Algorithm 2 in [84] is that MAARP uses Mirror Map while Algorithm 2 in [84] uses Euclidean projection. As already discussed in Subsection 2.1.3, the mirror constitutes a generalization of the Euclidean projection. Also discussed in Subsection 2.1.3, the use of a mirror map other than the Euclidean projection might result in a weaker dependence of the algorithm performance on the dimension of the strategy space (see Remark 12 and Section 4.8). Another difference between MAARP and Algorithm 2 in [84] is that the price update in Algorithm 2 in [84] requires two consecutive congestion states of a resource, i.e. $\mathbf{g}_r(\boldsymbol{X}_n)$ and $\mathbf{g}_r(\boldsymbol{X}_{n+1})$, while MAARP requires only $\mathbf{g}_r(\boldsymbol{X}_n)$. Moreover, in contrast to Algorithm 2 in [84], the price update $\boldsymbol{\Lambda}_{n+1}^r$ of MAARP can be done parallelly with the population update $\boldsymbol{X}_n$ since it does not depend on $\boldsymbol{X}_{n+1}$.

4.4. Convergence Analysis of MAARP

In this section, we investigate the convergence of primal iterate of the MAARP to a generalized Nash equilibrium of the game of our interest. The corresponding convergence result implies in particular that our pricing method can ensure that the strategy of the non-cooperative online learning agents' is sustainable w.r.t. the coupled (resource) constraints.

In general, the generalized Nash equilibrium (GNE) is not unique. Therefore, we need to specify the GNE of our interest: We introduce the notion of the so-called variational Nash equilibrium of the NGCC $\mathit{\Gamma}$ with the constraints $\mathcal{C}$ constituting a subclass of generalized Nash equilibrium arises naturally with the dynamic structure of MDAL. This subclass of generalized Nash equilibrium arises naturally with the first-order dynamic (w.r.t. $\mathbf{v}$) structure of MDAL as it is defined the solution $\mathrm{SOL}(\mathcal{Q}, \mathbf{v})$ of the well-known variational inequality $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ (See Definition 2.4). To improve the reading flow, we defer the detailed discussion on those aspects to the appendix.

In order to keep the argumentation short by avoiding the notion of convergence of a sequence to a set, we consider the case where the solution of $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ (and respectively $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}}))$ is unique. This property holds as asserted by Proposition 4.17 if $\mathbf{v}$ is strictly monotone, which holds if the utility function strictly convex in the sense that:

$$\mathrm{u}^{(i)}((\cdot), \boldsymbol{x}^{(-i)}) \text{ strictly convex } \quad \forall i \in [N],\ \boldsymbol{x}^{(-i)} \in \mathcal{X}_{-i}$$

For the analysis in this section, we use the notations:

$$\boldsymbol{M}_n := (\boldsymbol{M}_n^{(i)})_i \quad \text{and} \quad \sigma_n^2 := \mathbb{E}[\|\boldsymbol{M}_n\|_*^2].$$

We have the following convergence statement for the iterate of MAARP:

Theorem 4.1: *Suppose that the Assumption 4.3 holds. Moreover suppose that* $\mathbf{v}$ *is in addition strictly monotone. Let the augmentation function* θ_n *is given by* (4.5). *If the step size sequence* $(\gamma_n)_n$ *satisfies the properties:*

$$\sum_{k=0}^{\infty} \gamma_k = \infty, \qquad \sum_{k=0}^{\infty} \gamma_k^2 < \infty \tag{4.6}$$

$$\sum_{k=0}^{\infty} \gamma_k^2 \sigma_{k+1}^2 < \infty \tag{4.7}$$

and if (α_n) *and* (γ_n) *satisfy:*

$$\sum_{k=0}^{\infty} \gamma_k \alpha_k < \infty \tag{4.8}$$

and for large enough $k \in \mathbb{N}_0$*:*

$$\gamma_k \left(\alpha_k^2 + \frac{C_1^2}{K} \right) - \frac{\alpha_k}{2} \leqslant 0, \tag{4.9}$$

then the primal iterate $(\boldsymbol{X}_n)$ *of MAARP converges to the unique variational Nash equilibrium* $x_* = \mathrm{SOL}(\mathcal{Q}, \mathbf{v})$.

Before we provide proof of the above statement, we give in the following some examples of step size sequences that fulfill the assumptions of the above Theorem:

Remark 7: Assume that the noise is persistent, i.e. there exists $\sigma > 0$ s.t. $\sigma_k \leqslant \sigma$ for all $k \in \mathbb{N}$. The condition (4.7) turns to:

$$\sum_{k=0}^{\infty} \gamma_k^2 < \infty.$$

Consequently we can eliminate the redundant condition (4.7). (4.7) includes the possibilities that:

$$\gamma_n = \Theta(n^{-p}), \quad p \in (1/2, 1],$$

but rules out the possibilities that:

$$\gamma_n = \Theta(n^{-p}), \quad p \in [0, 1/2].$$

Now, let be:

$$\gamma_n = \frac{\gamma}{(n+1)^p}, \quad \text{where} \quad \gamma > 0 \quad \text{and} \quad p \in (1/2, 1] \text{ arbitrary.}$$

If we choose :

$$\alpha_n = \alpha\gamma_n, \quad \text{for an } \alpha > 0,$$

then (4.8) is fulfilled. Moreover, we have:

$$\gamma_n \left(\alpha_n^2 + \frac{C_1^2}{K} \right) - \frac{\alpha_n}{2} = \gamma_n \left[\alpha^2\gamma_n^2 + \frac{C_1^2}{K} - \frac{\alpha}{2} \right].$$

In case that:

$$\alpha > 2C_1^2/K,$$

we can find $c > 0$ such that:

$$\gamma_n(\alpha_n^2 + \frac{C_1^2}{K}) - \frac{\alpha_n}{2} \leqslant \gamma_n \left[\alpha^2\gamma_n^2 - c \right].$$

Since the R.H.S. of above inequality is negative for large n, (4.9) is fulfilled by the choice:

$$\alpha > 2\frac{C_1^2}{K}.$$

Remark 8: An immediate application of the Fubini's Theorem and the tower property yields that the condition (4.7) implies the condition that:

$$\sum_{k=0}^{\infty} \gamma_k^2 \mathbb{E}_k[\|\boldsymbol{M}_{k+1}\|_*^2] < \infty, \quad \text{a.s.}$$

which is often used in the martingale analysis. To see the implication, notice that we have by Fubini's Theorem and the tower property of the conditional expectation:

$$\infty > \sum_{k=0}^{\infty} \gamma_k^2 \sigma_{k+1}^2 = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma_k^2 \mathbb{E}_k[\|\boldsymbol{M}_{k+1}\|_*^2]\right].$$

Since the infinite sum within the expectation is non-negative, we have the desired statement.

In order to differentiate our approach to one of the nearest prior works, we give the following remarks:

Remark 9 (Relation to Stochastic Approximation Theory): At first sight, one may think that the dynamic (4.4) is an instance of the stochastic approximation algorithm (see e.g. [112–114]) having the archetypical form:

$$\boldsymbol{Z}_{n+1} = \boldsymbol{Z}_n + \gamma_n \left[h(\boldsymbol{Z}_n) + \boldsymbol{M}_{n+1}\right], \tag{4.10}$$

for a Lipschitz continuous vector field h. So, the question might arise whether one can immediately obtain Theorem 4.1 using stochastic approximation theory, whose approach ("ODE approach") consists of considering (4.10) as a Cauchy-Euler approximation of the ordinary differential equation (ODE) :

$$\dot{\boldsymbol{Z}}_t = h(\boldsymbol{Z}_t).$$

However, taking a detailed look at (4.4), one can recognize that our proposed algorithm differs from (4.10) by the non-linear mappings, i.e., $\boldsymbol{\Phi}$ and $\boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}}$, ensuring that the corresponding dynamic remains in the feasible sets. Moreover, in contrast to stochastic approximation theory, we do not require any Lipschitz condition on the vector field $\mathbf{v}$ needed, in order that the solution of the corresponding ODE uniquely exists. Even if we require the Lipschitz continuity $\mathbf{v}$, taking a similar ODE approach as done in the stochastic approximation theory by defining the ODE approximation of (4.4) as:

$$\begin{aligned} \boldsymbol{X}_t &= \boldsymbol{\Phi}(\boldsymbol{Y}_t), \quad \dot{\boldsymbol{Y}}_t = \mathbf{v}(\boldsymbol{X}_t) - [\nabla \mathbf{g}(\boldsymbol{X}_t)]^{\mathrm{T}} \boldsymbol{\Lambda}_t \\ \boldsymbol{\Lambda}_t &= \boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}}(\boldsymbol{\Gamma}_t), \quad \dot{\boldsymbol{\Gamma}}_t = \mathbf{g}(\boldsymbol{X}_t) - \nabla_{\boldsymbol{\Lambda}} \theta_t(\boldsymbol{\Lambda}_t). \end{aligned}$$

would surely require intricate argumentation: Starting with showing that the solution of above ODE uniquely exists – which is not immediately follows since a mirror map is in general not invertible. Nevertheless, the work [76] investigating the dynamic 4.3 constituting the fundament of our proposed algorithm (4.4) follows the intricate ODE approach by using techniques provided in [112]. We give a more detailed comment on this aspect in Remark 10.

Besides, the requirements (4.6) and (4.7) which are the usual summability condition in the stochastic approximation theory (see e.g. equation (2) in [113]) might lead some readers to think that Theorem 4.1 is an easy consequence of the stochastic approximation theory. However, in order to derive the convergence result, we need to postulate additional requirements: The requirement (4.8) that ensures the decreasing influence of the non-equilibrium of the prices, and the requirement (4.9) that ensures that the price update tracks the population dynamic.

Remark 10 (Relation to [76]): One may naively think that by defining a canonical new game with a new player controlling the dual variable analogous to the process done in Subsection 4.9.2, Theorem 4.1 is a simple assertion of Theorem 4.7 [76]. This claim is not valid since Theorem 4.7 in [76] relies on the fact that the constraint set of each player is compact and the constraint set $\mathbb{R}^R_{\geqslant 0}$ of the dual variable is unbounded. Furthermore, again by the fact that the constraint of the dual variable is not compact, we cannot imitate the approach done in [76] based on the theory provided in [112]. This hurdle motivates us to search for another way (Lemma 4.5) to generate the convergence statement (Theorem 4.1) from the recurrence result (Lemma 4.4). Our approach (Lemma 4.5) is much simpler than that given in [76] and can also be used to generate Theorem 4.7 in [76]. Another difference of our work to [76] is our weaker noise assumption. In [76], it is assumed that there exists $\sigma > 0$ s.t. $\mathbb{E}_n[\|\boldsymbol{M}_{n+1}\|_*^2] \leqslant \sigma^2$ a.s. for all $n \in \mathbb{N}_0$. Assuming (4.6), which is also done in [76], and applying Fubini's Theorem and the tower property, the latter observation implies (4.7).

4.4.1. Bound for Primal-Dual Iterate

The first step to prove Theorem 4.1 is to investigate the distance between primal-dual iterate of the MAARP to the solution of $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$. As a distance function, we use:

$$\tilde{\mathrm{F}}((\boldsymbol{x}, \boldsymbol{\lambda}), (\boldsymbol{\Phi}(\boldsymbol{y}), \tilde{\boldsymbol{\lambda}})) := \overline{\mathrm{F}}(\boldsymbol{x}, \boldsymbol{y}) + (\|\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}}\|_2^2/2)$$

where $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{y} \in \mathbb{R}^D$ and $\boldsymbol{\lambda} \in \mathbb{R}^R$. The following result gives a bound for:

$$\mathcal{V}_n(\boldsymbol{z}) := \tilde{\mathrm{F}}(\boldsymbol{z}, \boldsymbol{Z}_n) - \tilde{\mathrm{F}}(\boldsymbol{z}, \boldsymbol{Z}_0),$$

where:

$$\boldsymbol{z} := (\boldsymbol{x}, \boldsymbol{\lambda}) \in \mathcal{X} \times \mathbb{R}^R_{\geqslant 0} \quad \text{and} \quad \boldsymbol{Z}_n = (\boldsymbol{X}_n, \boldsymbol{\Lambda}_n)$$

Theorem 4.2: *Let $C_1, C_2, C_3 > 0$ be constants s.t. for all $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{\lambda} \in \mathbb{R}^R_{\geqslant 0}$ it holds:*

$$\|\nabla \mathbf{g}(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{\lambda}\|_* \leqslant C_1 \|\boldsymbol{\lambda}\|_2, \quad \|\mathbf{v}(\boldsymbol{x})\|_* \leqslant C_2, \quad \|\mathbf{g}(\boldsymbol{x})\|_2 \leqslant C_3, \tag{4.11}$$

and suppose that for all n, θ_n is continuously differentiable and K_n-strongly convex. For *all* $\boldsymbol{z} := (\boldsymbol{x}, \boldsymbol{\lambda}) \in \mathcal{X} \times \mathbb{R}^R_{\geqslant 0}$*, it holds:*

$$\mathcal{V}_n(\boldsymbol{z}) \leqslant \sum_{k=0}^{n-1} \gamma_k \eta_k(\boldsymbol{z}) + 2\left(\frac{C_2^2}{K} + \frac{C_3^2}{2}\right) \sum_{k=0}^{n-1} \gamma_k^2 + \sum_{k=0}^{n-1} \gamma_k \Psi_k(\boldsymbol{\Lambda}_k, \boldsymbol{\lambda}) + \boldsymbol{S}_n(\boldsymbol{x}) + \frac{2}{K}\boldsymbol{R}_n \tag{4.12}$$

where:

$$\begin{aligned}
\eta_k(\boldsymbol{z}) &:= \langle \boldsymbol{Z}_k - \boldsymbol{z}, \tilde{\mathrm{v}}(\boldsymbol{Z}_k) \rangle, \\
\boldsymbol{S}_n(\boldsymbol{x}) &:= \sum_{k=0}^{n-1} \gamma_k \tilde{\boldsymbol{M}}_k(\boldsymbol{x}), \;\; \tilde{\boldsymbol{M}}_k(\boldsymbol{x}) := \langle \boldsymbol{X}_k - \boldsymbol{x}, \boldsymbol{M}_{k+1} \rangle, \\
\boldsymbol{R}_n &:= \sum_{k=0}^{n-1} \gamma_k^2 \|\boldsymbol{M}_{k+1}\|_*^2, \\
\Psi_k(\boldsymbol{\Lambda}_k, \boldsymbol{\lambda}) &= \theta_k(\boldsymbol{\lambda}) - \frac{K_k}{2}\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2 + \left[\gamma_k \left(\|\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k)\|_2^2 + \frac{C_1^2}{K}\|\boldsymbol{\Lambda}_k\|_2^2\right) - \theta_k(\boldsymbol{\Lambda}_k)\right]
\end{aligned} \tag{4.13}$$

Proof: First, define:

$$\mathcal{E}_k^{(1)}(\boldsymbol{x}) := \overline{\mathrm{F}}(\boldsymbol{x}, \boldsymbol{Y}_k).$$

By inserting the iterate of the MAARP into the bound given in Proposition 2.4, we have:

$$\begin{aligned}
\mathcal{E}_{k+1}^{(1)}(\boldsymbol{x}) - \mathcal{E}_k^{(1)}(\boldsymbol{x}) \leqslant{}& \gamma_k \langle \boldsymbol{X}_k - \boldsymbol{x}, \mathbf{v}(\boldsymbol{X}_k) - [\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k \rangle + \gamma_k \langle \boldsymbol{X}_k - \boldsymbol{x}, \boldsymbol{M}_{k+1} \rangle \\
&+ \frac{\gamma_k^2}{2K} \|[\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k + \mathbf{v}(\boldsymbol{X}_k) + \boldsymbol{M}_{k+1}\|_*^2
\end{aligned}$$

Triangle inequality, the inequality $(\sum_{i=1}^K a_i)^2 \leqslant K \sum_{i=1}^K a_i^2$, and the assumption (4.11), gives:

$$\begin{aligned}
\|[\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k + \mathbf{v}(\boldsymbol{X}_k) + \boldsymbol{M}_{k+1}\|_*^2 &\leqslant \left(\|[\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k\|_* + \|\mathbf{v}(\boldsymbol{X}_k) + \boldsymbol{M}_{k+1}\|_*\right)^2 \\
&\leqslant 2\|[\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k\|_*^2 + 2\|\mathbf{v}(\boldsymbol{X}_k) + \boldsymbol{M}_{k+1}\|_*^2 \\
&\leqslant 2\underbrace{\|[\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k\|_*^2}_{\leqslant C_1^2 \|\boldsymbol{\Lambda}\|_2^2} + 2\left(2\underbrace{\|\mathbf{v}(\boldsymbol{X}_k)\|_*^2}_{\leqslant C_2^2} + 2\|\boldsymbol{M}_{k+1}\|_*^2\right)
\end{aligned} \tag{4.14}$$

Combining both computations, we have:

$$\begin{aligned}
\mathcal{E}_{k+1}^{(1)}(\boldsymbol{x}) - \mathcal{E}_k^{(1)}(\boldsymbol{x}) \leqslant{}& \gamma_k \langle \boldsymbol{X}_k - \boldsymbol{x}, \mathbf{v}(\boldsymbol{X}_k) - [\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k \rangle \\
&+ \gamma_k \tilde{\boldsymbol{M}}_{k+1} + \frac{\gamma_k^2}{K}\left(C_1^2 \|\boldsymbol{\Lambda}_k\|_2^2 + 2(C_2^2 + \|\boldsymbol{M}_{k+1}\|_*^2)\right)
\end{aligned}$$

Summing above inequality over all $k = 0, \ldots, n-1$ and subsequent telescoping, we obtain the following upper bound for $\mathcal{V}_n^{(1)}(\boldsymbol{x})$:

$$\begin{aligned}\mathcal{V}_n^{(1)}(\boldsymbol{x}) \leqslant & \sum_{k=0}^{n-1} \gamma_k \left[\langle \boldsymbol{X}_k - \boldsymbol{x}, \mathrm{v}(\boldsymbol{X}_k) - [\nabla \mathbf{g}(\boldsymbol{X}_k)]^T \boldsymbol{\Lambda}_k \rangle\right] \\ & + \sum_{k=0}^{n-1} \frac{\gamma_k^2 C_1^2}{K} \|\boldsymbol{\Lambda}_k\|_2^2 + \boldsymbol{S}_n(\boldsymbol{x}) + \boldsymbol{R}_n + \frac{2C_2^2}{K} \sum_{k=0}^{n-1} \gamma_k^2,\end{aligned}$$

where:

$$\mathcal{V}_n^{(1)}(\boldsymbol{x}) := \overline{\mathrm{F}}(\boldsymbol{x}, \boldsymbol{Y}_n) - \overline{\mathrm{F}}(\boldsymbol{x}, \boldsymbol{Y}_0).$$

Now we want to bound:

$$\mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) := \frac{\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2 - \|\boldsymbol{\Lambda}_0 - \boldsymbol{\lambda}\|_2^2}{2}.$$

For that sake, we compute:

$$\|\boldsymbol{\Lambda}_{k+1} - \boldsymbol{\lambda}\|_2^2 = \|\boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}} \left(\boldsymbol{\Lambda}_k - \gamma_k \left(\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k) - \mathbf{g}_k(\boldsymbol{X}_k)\right)\right) - \boldsymbol{\lambda}\|_2^2 \tag{4.15}$$

$$\leqslant \|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda} - \gamma_k \left(\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k) - \mathbf{g}_k(\boldsymbol{X}_k)\right)\|_2^2 \tag{4.16}$$

$$= \|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2 - 2\gamma_k \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, -\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k) - \mathbf{g}_k(\boldsymbol{X}_k)\rangle + \gamma_k^2 \|\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k) - \mathbf{g}_k(\boldsymbol{X}_k)\|_2^2 \tag{4.17}$$

where (4.15) follows by setting the update (4.1), (4.16) from the non-expansive property of the Euclidean projection, i.e.:

$$\|\boldsymbol{\Phi}(\boldsymbol{y}) - \boldsymbol{x}\|_2 \leqslant \|\boldsymbol{y} - \boldsymbol{x}\|_2, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^R,$$

and (4.17) from the identity:

$$\|\boldsymbol{x} + \boldsymbol{y}\|_2^2 = \|\boldsymbol{x}\|_2^2 + 2\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \|\boldsymbol{y}\|_2^2.$$

By similar computation as in (4.14), we have:

$$\|\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k) - \mathbf{g}_k(\boldsymbol{X}_k)\|_2^2 \leqslant 2(C_3^2 + \|\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k)\|_2^2) \tag{4.18}$$

Combining (4.17) and (4.18), we obtain:

$$\begin{aligned}\mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) \leqslant & \sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{g}(\boldsymbol{X}_k)\rangle - \sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k)\rangle \\ & + \sum_{k=0}^{n-1} \gamma_k^2 (C_3^2 + \|\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k)\|_2^2).\end{aligned}$$

Now, since θ_k is K_k-strongly convex for all k, we have:

$$\langle \boldsymbol{\lambda} - \boldsymbol{\Lambda}_k, \nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k) \rangle \leqslant \theta_k(\boldsymbol{\lambda}) - \theta_k(\boldsymbol{\Lambda}_k) - \frac{K_k}{2} \|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2,$$

and consequently:

$$\begin{aligned} \mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) \leqslant \sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{g}(\boldsymbol{X}_k) \rangle + \sum_{k=0}^{n-1} \gamma_k [\theta_k(\boldsymbol{\lambda}) - \theta_k(\boldsymbol{\Lambda}_k)] \quad (4.19) \\ - \sum_{k=0}^{n-1} \frac{\gamma_k K_k}{2} \|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2 + \sum_{k=0}^{n-1} \gamma_k^2 (C_3^2 + \|\nabla_{\boldsymbol{\Lambda}_k} \theta_k(\boldsymbol{\Lambda}_k)\|_2^2). \quad (4.20) \end{aligned}$$

Combining the bounds for $\mathcal{V}_n^{(1)}(\boldsymbol{x})$ and $\mathcal{V}_n^{(2)}(\boldsymbol{\lambda})$, we obtain the desired bound for:

$$\mathcal{V}_n(\boldsymbol{x}, \boldsymbol{\lambda}) = \mathcal{V}_n^{(1)}(\boldsymbol{x}) + \mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) \qquad \blacksquare$$

Regarding the estimate given in Theorem 4.2, one possibility to gain control over $\mathcal{V}_n$ is to eliminate the dependency of the upper bound (4.12) on $\boldsymbol{\Lambda}_k$, $k \in [n-1]_0$. We accomplish this goal by choosing the augmentation functions as in (4.5), each for all $n \in \mathbb{N}_0$ a α_n-strongly convex augmentation function, yielding:

$$\begin{aligned} 2\Psi_n(\boldsymbol{\Lambda}_n, \boldsymbol{\lambda}) &= \alpha_n \|\boldsymbol{\lambda}\|_2^2 + \beta_n \|\boldsymbol{\Lambda}_n\|_2^2 - \alpha_k \|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2 \\ \frac{\beta_n}{2} &:= \left[\gamma_n \left(\alpha_n^2 + \frac{C_1^2}{K} \right) - \frac{\alpha_n}{2} \right]. \end{aligned} \quad (4.21)$$

and subsequently ensuring $\beta_k \leqslant 0$ by appropriately choosing α_k and γ_k (see (4.9)).

4.4.2. Control over Noise

In order to gain control over the disturbance due to the noise in the first-order feedback represented by the sums $\boldsymbol{S}_n(\boldsymbol{x})$ and $\boldsymbol{R}_n$ occurring in the upper bound (4.12), we apply the Doob's martingale convergence Theorem known in the literature of martingale theory (see e.g., [56]):

Lemma 4.3: *Suppose that* $(\boldsymbol{M}_n)$ *is a square integrable* $\mathbb{R}^D$*-valued* $\mathbb{F}$*-martingale difference sequence and that* (4.7) *holds. Then for the partial sums* $\boldsymbol{S}_n = \boldsymbol{S}_n(\boldsymbol{x})$ *and* $\boldsymbol{R}_n = \boldsymbol{R}_n(\boldsymbol{x})$, $n \geqslant 0$, *given in* (4.13), *it holds:*

1. $(\boldsymbol{S}_n)_n$ *is a mean zero square integrable* $\mathbb{F}$*-martingale and* $(\boldsymbol{R}_n)_n$ *is a non-negative* $\mathbb{F}$*-sub-martingale*

2. *There exists a square integrable real R.V.* R_∞ *and an integrable real RV s.t.* $(\boldsymbol{S}_n)_n$ *converges a.s. and in* L^2 *to* S_∞ *and* $(\boldsymbol{R}_n)_n$ *converges a.s. and in* L^1 *to* R_∞

Proof: It holds since $\boldsymbol{M}_k$ is square integrable for all k:

$$\mathbb{E}[|\boldsymbol{R}_n|] = \sum_{k=0}^{n-1} \gamma_k^2 \mathbb{E}[\|\boldsymbol{M}_{k+1}\|_*^2] < \infty.$$

Moreover by the triangle inequality and the Hölder's inequality, it yields:

$$|\boldsymbol{S}_n|^2 \leqslant n \sum_{k=0}^{n-1} \gamma_t^2 \left|\tilde{\boldsymbol{M}}_k\right|^2 \leqslant n \sum_{k=0}^{n-1} \gamma_k^2 \|X_k - \boldsymbol{x}\|^2 \|\boldsymbol{M}_{k+1}\|_*^2 \leqslant T C_{\mathcal{X}}^2 \boldsymbol{R}_n,$$

where $C_{\mathcal{X}} > 0$ is a constant, whose existence is ensured by compactness of $\mathcal{X}$. This observation and the integrability of $\boldsymbol{R}_n$ give the square integrability of $\boldsymbol{S}_n$.

Now, since $\|\boldsymbol{M}_{k+1}\|^2$ is non-negative, it follows immediately that:

$$\mathbb{E}_{n-1}[\boldsymbol{R}_n] \geqslant \boldsymbol{R}_{n-1}.$$

This observation and the apparent fact that $(\boldsymbol{R}_n)$ is $\mathbb{F}$-adapted give that $(\boldsymbol{R}_n)$ is a sub-martingale.

Next, we show that $(\boldsymbol{S}_n)$ is a martingale. Since for all $k \in \mathbb{N}_0$, X_k is a measureable function of $(\boldsymbol{M}_\tau)_{\tau \leqslant k}$, it follows that $(\boldsymbol{X}_n)$ is adapted to $\mathbb{F}$. Consequently $(\boldsymbol{S}_n)_n$ is adapted to $\mathbb{F}$. Moreover the fact that $\boldsymbol{X}_n$ is $\mathcal{F}_n$-measureable asserts:

$$\mathbb{E}_k[\tilde{\boldsymbol{M}}_k] = \langle \boldsymbol{X}_k - \boldsymbol{x}, \mathbb{E}_k[\boldsymbol{M}_{k+1}] \rangle = 0. \tag{4.22}$$

So, consequently we have as desired $\mathbb{E}_n[\boldsymbol{S}_{n+1}] = \boldsymbol{S}_n$. It is immediate to see that this fact and (4.22) assert that $\mathbb{E}[\boldsymbol{S}_n] = 0$.

Now, we proof the statement 2). For $k, \tilde{k} \in \mathbb{N}_0$ with $k < \tilde{k}$, we have by tower property, by the fact that $\tilde{\boldsymbol{M}}_k$ is $\mathcal{F}_{\tilde{k}}$-measureable, and by (4.22):

$$\mathbb{E}[\tilde{\boldsymbol{M}}_k \tilde{\boldsymbol{M}}_{\tilde{k}}] = \mathbb{E}[\mathbb{E}_{\tilde{k}-1}[\tilde{\boldsymbol{M}}_k \tilde{\boldsymbol{M}}_{\tilde{k}}]] = \mathbb{E}[\tilde{\boldsymbol{M}}_k \mathbb{E}_k[\tilde{\boldsymbol{M}}_{\tilde{k}}]] = 0.$$

This asserts that:

$$\mathbb{E}[|\boldsymbol{S}_n|^2] = \sum_{k=0}^{n-1} \gamma_k^2 \mathbb{E}[|\tilde{\boldsymbol{M}}_k|^2] + \sum_{k \neq \tilde{k}} \gamma_k \gamma_{\tilde{k}} \mathbb{E}[\tilde{\boldsymbol{M}}_k \tilde{\boldsymbol{M}}_{\tilde{k}}] \leqslant C_{\mathcal{X}}^2 \sum_{k=0}^{n-1} \gamma_k^2 \mathbb{E}[\|\boldsymbol{M}_k\|_*^2].$$

Assumption (4.7) asserts that $\sum_{k=0}^{\infty} \gamma_k^2 \mathbb{E}[\|\boldsymbol{M}_{k+1}\|_*^2] < \infty$. So this observation and above inequality assert that $(|\boldsymbol{S}_n|^2)$ is uniformly convergence. Therefore by martingale convergence Theorem, we obtain the desired statement for $(|\boldsymbol{S}_n|^2)$. The fact that $(\boldsymbol{R}_n)$ is uniformly integrable follows immediately from (4.7), and the desired statement is a direct application of martingale convergence Theorem. ■

4.4.3. Convergence Proof

To show the Theorem 4.1, we use the Proposition 4.16 given in the Appendix. This proposition asserts that in order to show the convergence of the primal iterate to the variational Nash equilibrium of the NGCC Γ with the constraints $\mathcal{C}$, we can instead investigate the convergence behavior of both the population - and the price iterate to the solution of the extended variational inequality $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$ where $\tilde{\mathbf{v}}$ is the KKT operator corresponding to $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ given in (4.54). A detailed discussion on this aspect is given in the Subsubsection 4.9.2 in the Appendix.

A first step toward this direction is to combine the results obtained in the preceding subsections and show the recurrence of $\boldsymbol{Z}_n$ around the solution of the variational inequality $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$:

Lemma 4.4: *Suppose that the Assumptions 3.3, 4.1, and 4.2 hold, and that* $\mathbf{v}$ *is strictly monotone. Let the augmentation function* θ_t *is given by* (4.5)*. If* $(\gamma_n)_{n\in\mathbb{N}_0}$*,* $(\boldsymbol{M}_n)_{n\in\mathbb{N}}$*, and* $(\alpha_n)_{n\in\mathbb{N}_0}$ *fulfill* (4.7), (4.6), (4.8)*, and* (4.9)*, then the primal-dual iterate* $(\boldsymbol{Z}_n)_n$ *a.s. recurs in all the neighbors (w.r.t.* $\|\cdot\|_*$*) of the unique variational Nash equilibrium* $\boldsymbol{z}_* \in \mathrm{SOL}(\mathcal{X}\times\mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$ *of* Γ*, i.e. a.s. there exists a subsequence* $(\boldsymbol{Z}_{n_k})_k$ *of* $(\boldsymbol{Z}_n)_n$ *which converges to* $\boldsymbol{z}_*$ *w.r.t.* $\|\cdot\|_*$.

Proof: Clearly since $\boldsymbol{X}_n \in \mathcal{X}$ for all $n \in \mathbb{N}$ with $\mathcal{X}$ compact, the sequence $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$ is non-explosive, i.e. $\|\boldsymbol{X}_n\| < \infty$ for all $n \in \mathbb{N}$. Moreover via standard Grönwall's argumentation, it follows that $\boldsymbol{\Lambda}_n$ is also non-explosive (see Subsection 4.9.3). Those facts assert that we can assume without loss of generality that (4.9) holds for all $k \in \mathbb{N}$. Notice that by Theorem 4.2, it follows that:

$$\mathcal{V}_n = \mathcal{V}_n(\boldsymbol{z}_*) \leqslant \tau_n \left(\frac{\sum_{k=0}^{n-1} \gamma_k \eta_k}{\tau_n} + \frac{\tilde{C}_1 \sum_{k=0}^{n-1} \gamma_k^2}{\tau_n} + \frac{\sum_{k=0}^{n-1} \gamma_k \Psi_k}{\tau_n} + \frac{\boldsymbol{S}_n + \frac{2}{K}\boldsymbol{R}_n}{\tau_n} \right) \tag{4.23}$$

where :

$$\tau_n := \sum_{k=0}^{n-1} \gamma_k, \quad \Psi_k := \Psi_k(\boldsymbol{\Lambda}_k, \quad \boldsymbol{\lambda}_*), \quad \eta_k := \eta_k(\boldsymbol{z}^*), \text{and} \quad \tilde{C}_1 = 2((C_2^2/K) + (C_3^2/2)).$$

By our choice of the augmentation function and the condition (4.9), it follows that (see (4.5)):

$$\Psi_k \leqslant \alpha_k \|\boldsymbol{\lambda}_*\|_2^2/2.$$

Setting this estimate into (4.23), it follows that:

$$\frac{\mathcal{V}_n}{\tau_n} \leqslant \frac{\sum_{k=n_0}^{n-1} \gamma_k \eta_k}{\tau_n} + \frac{\tilde{C}_1 \sum_{k=n_0}^{n-1} \gamma_k^2}{\tau_n} + \frac{\|\boldsymbol{\lambda}_*\|_2^2 \sum_{k=n_0}^{n-1} \gamma_k \alpha_k}{2\tau_n} + \frac{\boldsymbol{S}_n + \frac{2}{K}\boldsymbol{R}_n}{\tau_n}. \tag{4.24}$$

Now, Lemma 4.3 asserts that both $(\boldsymbol{S}_n)_n$ and $(\boldsymbol{R}_n)_n$ converge a.s. and therefore a.s. there exists a constant $A > 0$ s.t.:

$$\boldsymbol{S}_n + \frac{2}{K}\boldsymbol{R}_n \leqslant A, \quad \forall n.$$

This and (4.24) yield that:

$$\frac{\mathcal{V}_n}{\tau_n} \leqslant \frac{\sum_{k=n_0}^{n-1}\gamma_k\eta_k}{\tau_n} + \frac{\tilde{C}_1\sum_{k=n_0}^{n-1}\gamma_k^2}{\tau_n} + \frac{\|\boldsymbol{\lambda}_*\|_2^2\sum_{k=n_0}^{n-1}\gamma_k\alpha_k}{2\tau_n} + \frac{A}{\tau_n}. \tag{4.25}$$

Let U be an arbitrary neighborhood of $\boldsymbol{z}_*$ w.r.t. the norm on $\mathcal{V} \times \mathbb{R}^R$. Suppose that $\boldsymbol{Z}_n \notin U$ for all sufficiently large $n \geqslant 0$. W.l.o.g. we assume that $\boldsymbol{Z}_n \notin U$ for all $n \geqslant 0$. Since $\mathbf{v}$ is strictly monotone and and therefore also $\tilde{\mathbf{v}}$ strictly monotone, it follows from (2.6) that we can find $c > 0$ s.t.:

$$\eta_k \leqslant -c, \quad \forall k \geqslant 0.$$

This yields:

$$\mathcal{V}_n \leqslant \tau_n\left(-c + \tilde{C}_1\frac{\sum_{k=n_0}^{n-1}\gamma_k^2}{\tau_n} + \frac{\sum_{k=n_0}^{n-1}\gamma_k\alpha_k}{2\tau_n}\|\boldsymbol{\lambda}_*\|_2^2 + \frac{A}{\tau_n}\right). \tag{4.26}$$

Now, (4.6) and (4.8) asserts that:

$$\tilde{C}_1\frac{\sum_{k=n_0}^{n-1}\gamma_k^2}{\tau_n} + \frac{\|\boldsymbol{\lambda}_*\|_2^2\sum_{k=n_0}^{n-1}\gamma_k\alpha_k}{2\tau_n} + \frac{A}{\tau_n} \to 0, \quad \text{as } n \to \infty. \tag{4.27}$$

Finally, since:

$$\tau_n \to \infty, \quad \text{as} \quad n \to \infty,$$

we have from (4.27) and (4.26) that:

$$\mathcal{V}_n \to -\infty, \quad n \to \infty \quad \text{a.s.},$$

which contradicts with the fact that:

$$\mathcal{V}_n \geqslant -\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_0) > -\infty.$$

Thus a.s. $\boldsymbol{Z}_n \in U$ for infinitely many $n \geqslant 0$. ■

Remark 11: Taking a glance into the proof of Lemma 4.4, one notices that in case that there is no noise, the requirements (4.6) and (4.8) in this lemma can be replaced by the weaker requirements:

$$\sum_{k=0}^{\infty}\gamma_k = \infty, \quad \frac{\sum_{k=1}^{n}\gamma_k^2}{\sum_{k=1}^{n}\gamma_k} \to 0, \quad \frac{\sum_{k=1}^{n}\gamma_k\alpha_k}{\sum_{k=1}^{n}\gamma_k} \to 0, \quad n \to \infty,$$

The final auxiliary step, we need for showing Theorem 4.1, is the following result in spirit of Robbin-Siegmund's Lemma (see e.g. Therom 1.3.12 in [115]):

Lemma 4.5: *Suppose that the Assumptions 3.3, 4.1, 4.2, and 4.3 hold, and that* $(\gamma_n)_{n\in\mathbb{N}_0}$*,* $(\boldsymbol{M}_n)_{n\in\mathbb{N}}$*, and* $(\alpha_n)_{n\in\mathbb{N}_0}$ *fulfill* (4.7) (4.6), (4.8), *and* (4.9). *Let* $\boldsymbol{z}_*$ *be a variational Nash equilibrium, i.e. solution of* $\boldsymbol{z}_* \in \mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$*. Then a.s.* $(\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n))_n$ *converges to a finite RV.*

Proof: W.l.o.g. we can assume that (4.9) holds for all n. We have by (4.24):

$$\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_{n+1}) \leqslant \tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) + \gamma_n\eta_n + \frac{\gamma_n\alpha_n\|\boldsymbol{\lambda}_*\|_2^2}{2} + \gamma_n^2\tilde{C}_1 + \gamma_n\tilde{\boldsymbol{M}}_n + \gamma_n^2\|\boldsymbol{M}_{n+1}\|_*^2. \tag{4.28}$$

Now, since $\boldsymbol{Z}_n$ is $\mathcal{F}_n$ measureable and $\eta_n = \eta_n(\boldsymbol{z}^*) \leqslant 0$ for all n (This follows from the assumptions that $\boldsymbol{z}^*$ is the solution of $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$ and that $\tilde{\mathbf{v}}$ is monotone) and since $\mathbb{E}_n[\tilde{\boldsymbol{M}}_n] = 0$, we have by taking the expectation of (4.28) given $\mathcal{F}_n$:

$$\mathbb{E}[\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_{n+1})|\mathrm{F}_n] \leqslant \tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) + \zeta_n, \tag{4.29}$$

where:

$$\zeta_n := \frac{\gamma_n\alpha_n\|\boldsymbol{\lambda}_*\|_2^2}{2} + \gamma_n^2\tilde{C}_1 + \gamma_n^2\mathbb{E}[\|\boldsymbol{M}_{n+1}\|_*^2|\mathcal{F}_n].$$

Furthermore, notice that by the assumptions (4.6), (4.8), (4.7):

$$(\zeta_n)_n \text{ is a non-negative } \mathbb{F}\text{-adapted sequence satisfying } \sum_{k=0}^{\infty}\mathbb{E}[\zeta_k] < \infty \text{ a.s.} \tag{4.30}$$

Indeed, this holds by the tower property of conditional expectation:

$$\sum_{k=0}^{\infty}\mathbb{E}\left[\gamma_k^2\mathbb{E}[\|\boldsymbol{M}_{k+1}\|_*^2]\right] = \sum_{k=0}^{\infty}\mathbb{E}\left[\gamma_k^2\mathbb{E}[\|\boldsymbol{M}_{k+1}\|_*^2]\right] = \sum_{k=0}^{\infty}\gamma_k^2\sigma_k^2,$$

by the Fubini's Theorem and the tower property of conditional expectation.

Now, consider the sequence:

$$Y_n := \tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) - \sum_{k=0}^{n-1}\zeta_k.$$

Clearly (Y_n) is $\mathbb{F}$-adapted and integrable. The fact that (ζ_n) is $\mathbb{F}$-adapted (see (4.30)) gives:

$$\mathbb{E}_n[Y_{n+1}] = \mathbb{E}_n[\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_{n+1})] - \mathbb{E}_n\left[\sum_{k=0}^{n}\zeta_k\right] = \mathbb{E}_n[\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_{n+1})] - \sum_{k=0}^{n}\zeta_k.$$

So combining this and (4.29), we have:

$$\mathbb{E}_n[Y_{n+1}] \leqslant \tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) + \zeta_n - \sum_{k=0}^{n}\zeta_k = \tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) - \sum_{k=0}^{n-1}\zeta_k = Y_n,$$

and thus we have that:

$$(Y_n) \quad \text{is a supermartingale.} \tag{4.31}$$

Moreover for all $n \geqslant 0$ we have since $\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) \geqslant 0$

$$\mathbb{E}[Y_n] \geqslant -\mathbb{E}\left[\sum_{k=0}^{n} \zeta_k\right] = -\sum_{k=0}^{n} \mathbb{E}\left[\zeta_k\right] \geqslant -\sum_{k=0}^{\infty} \mathbb{E}\left[\zeta_k\right] > -\infty, \quad \forall n \in \mathbb{N},$$

and consequently:

$$\sup_{n\in\mathbb{N}} \mathbb{E}[[Y_n]_-] < \infty. \tag{4.32}$$

Both (4.31) and (4.32), and the Doob's martingale convergence Theorem (Theorem 2.8) yield the a.s. convergence of (Y_n). Finally, since:

$$0 \leqslant \sum_{k=0}^{\infty} \zeta_k < \infty \quad \text{a.s.,}$$

which is a consequence of $\sum_{k=0}^{\infty} \mathbb{E}[\zeta_k] < \infty$ (c.f. Remark 8) it follows that $(\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n))_n$ converges a.s., as desired. ∎

Proof (Proof of Theorem 4.1): Let $\epsilon > 0$ and $\boldsymbol{z}_*$ the unique solution of $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$. Lemma 4.4 asserts (a.s.) the existence of a subsequence $(\boldsymbol{Z}_{n_k})_k$ of $(\boldsymbol{Z}_n)_n$ satisfying:

$$\boldsymbol{Z}_{n_k} \to z_*, \quad \text{as} \quad k \to \infty.$$

Consequently, by the reprocity condition we have that:

$$\mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_{n_k}) \to 0, \quad \text{as} \quad k \to \infty.$$

Furthermore, Lemma 4.5 asserts that $(\mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_n))_n$ converges a.s. to a finite random variable and therefore we have:

$$(\mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_n))_n \quad \text{is a.s. a Cauchy sequence.}$$

Consequently, a.s., we can choose n_0 large enough s.t. for all $n_k, n, \tilde{n} \geqslant n_0$:

$$\mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_{n_k}) \leqslant \epsilon/2 \quad \text{and} \quad \mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_{\tilde{n}}) \leqslant \mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_n) + \epsilon/2.$$

So for the latter event, we have for $n, n_k \geqslant n_0$:

$$\mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_n) \leqslant \mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_{n_k}) + \epsilon/2 \leqslant \epsilon.$$

Thus a.s. $\mathrm{F}(\boldsymbol{z}_*, \boldsymbol{Z}_n) \to 0$ as $n \to \infty$ and Proposition 2.4 asserts that $(\boldsymbol{Z}_n)$ converges a.s. to $\boldsymbol{z}_* = (\boldsymbol{x}_*, \boldsymbol{\lambda}_*)$ as $n \to \infty$. For the desired statement, notice finally that by Proposition

4.17, $\boldsymbol{x}^*$ is the unique variational Nash equilibrium. ■

4.5. Resource Constraint Violation Analysis

Theorem 4.1 asserts that in the long term and under suitable choices of algorithm parameters, the population iterate a.s. satisfies the coupled resource constraint. However, the guarantee is purely asymptotic. So, we aim in this section to provide a non-asymptotic guarantee for the decay of the amount of resource constraint violation of MAARP. To make the corresponding result more accessible, we assume in this section for convenience the following:

Assumption 4.4: *For all $n \in \mathbb{N}_0$, $\alpha_n = \alpha\gamma_n$ for an $\alpha > 0$. The trackability condition* (4.9) *holds for all $n \in \mathbb{N}$, which is in this context:*

$$\alpha^2\gamma_n^2 - \frac{\alpha}{2} + \frac{C_1^2}{K} \leqslant 0, \quad \forall n \in \mathbb{N}_0. \tag{4.33}$$

The noise is persistent in the sense that there exists $\sigma > 0$ s.t. $\sigma_k \leqslant \sigma$ for all $k \in \mathbb{N}$.

Vital for the derivation of a non-asymptotic guarantee is the following quantity:

$$\mathrm{CVio}_n^r := \frac{\sum_{k=0}^{n-1} \gamma_k \mathbf{g}_r(\boldsymbol{X}_k)}{\sum_{k=0}^{n-1} \gamma_k}.$$

Since $\mathbf{g}_r$ is convex for every $r \in [R]$, it follows from Jensen's inequality that CVio_n^r is an upper bound of the constraint violation on the resource r caused by the ergodic average of the population iterate at time $n-1$, i.e.:

$$\mathbf{g}_r(\overline{\boldsymbol{X}}_n) \leqslant \mathrm{CVio}_n^r, \quad \text{where } \overline{\boldsymbol{X}}_n := \frac{\sum_{k=0}^{n-1} \gamma_k \boldsymbol{X}_k}{\sum_{k=0}^{n-1} \gamma_k} \tag{4.34}$$

We start with the following auxiliary statement, which asserts that one can deduce the information about the load from the prices:

Lemma 4.6: *Suppose that :*

$$\boldsymbol{\Lambda}_0 = 0.$$

For all $r \in [R]$ and $n \in \mathbb{N}$:

$$\mathrm{CVio}_n^r \leqslant \frac{\|\boldsymbol{\Lambda}_n\|_2 + \sum_{k=0}^{n-1} \gamma_k \alpha_k \|\boldsymbol{\Lambda}_k\|_2}{\sum_{k=0}^{n-1} \gamma_k}$$

Proof: The definition of our price policy gives:

$$\boldsymbol{\Lambda}_{k+1}^r \geqslant \boldsymbol{\Lambda}_k^r + \gamma_k \mathbf{g}_r(\boldsymbol{X}_k) - \gamma_k \alpha_k \boldsymbol{\Lambda}_k^r.$$

So by summing this inequality and subsequent telescoping, we have since $\boldsymbol{\Lambda}_0 = 0$:

$$\sum_{k=0}^{n-1} \gamma_k \mathbf{g}_r(\boldsymbol{X}_k) \leqslant \boldsymbol{\Lambda}_n^r + \sum_{k=0}^{n-1} \gamma_k \alpha_k \boldsymbol{\Lambda}_k^r.$$

Now, this inequality and the fact that $\boldsymbol{\Lambda}_k^r \geqslant 0$ give:

$$\boldsymbol{\Lambda}_k^r \leqslant \sqrt{\sum_{r=1}^{R} (\boldsymbol{\Lambda}_k^r)^2} = \|\boldsymbol{\Lambda}_k\|_2.$$

Therefore, we have as desired:

$$\sum_{k=0}^{n-1} \gamma_k \mathbf{g}_r(\boldsymbol{X}_k) \leqslant \|\boldsymbol{\Lambda}_n\|_2 + \sum_{k=0}^{n-1} \gamma_k \alpha_k \|\boldsymbol{\Lambda}_k\|_2. \qquad \blacksquare$$

Lemma 4.7: *Suppose that:*

$$\boldsymbol{\Lambda}_0 = 0 \quad \textit{and} \quad \mathbf{Y}_0 = 0.$$

Then under Assumption 4.4, it holds:

$$\mathbb{E}[\|\boldsymbol{\Lambda}_n\|_2] \leqslant \mathcal{O}\left((1+\sigma) \sqrt{\sum_{k=0}^{n-1} \gamma_k^2} \right)$$

Proof: By taking the expectation of (4.29), and subsequent summing and telescoping, we can conclude that:

$$\mathbb{E}\left[\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_n) \right] \leqslant \tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_0) + \frac{\sum_{k=0}^{n-1} \gamma_k \alpha_k \|\boldsymbol{\lambda}_*\|_2^2}{2} + \sum_{k=0}^{n-1} \gamma_k^2 \tilde{C}_1 + \sum_{k=0}^{n-1} \gamma_k^2 \sigma_{k+1}^2.$$

Consequently since $\boldsymbol{Z}_0 = 0$:

$$\mathbb{E}\left[\frac{\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}_*\|^2}{2} \right] \leqslant \Delta\psi + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} + \frac{\sum_{k=0}^{n-1} \gamma_k \alpha_k \|\boldsymbol{\lambda}_*\|_2^2}{2} + \sum_{k=0}^{n-1} \gamma_k^2 \tilde{C}_1 + \sum_{k=0}^{n-1} \gamma_k^2 \sigma_{k+1}^2,$$

where:

$$\Delta\psi = \sum_{i=1}^{N} \left[\max_{\boldsymbol{x}^{(i)} \in \mathcal{X}_i} \psi_i(\boldsymbol{x}^{(i)}) - \min_{\boldsymbol{x}^{(i)} \in \mathcal{X}_i} \psi_i(\boldsymbol{x}^{(i)}) \right]. \tag{4.35}$$

Thus by Jensen's inequality and triangle inequality:

$$(\mathbb{E}[\|\boldsymbol{\Lambda}_n\|_2] - \|\boldsymbol{\lambda}_*\|_2)^2 \leqslant 2\Delta\psi + \|\boldsymbol{\lambda}_*\|_2^2 + \sum_{k=0}^{n-1} \gamma_k \alpha_k \|\boldsymbol{\lambda}_*\|_2^2 + 2\sum_{k=0}^{n-1} \gamma_k^2 \tilde{C}_1 + 2\sum_{k=0}^{n-1} \gamma_k^2 \sigma_{k+1}^2.$$

The desired statement follows from above since $\alpha_k = \alpha\gamma_k$ and $\sigma_k \leqslant \sigma$ for all k. $\blacksquare$

If $\gamma_k = \Theta(1/\sqrt{k})$, we have $\mathbb{E}[\|\boldsymbol{\Lambda}_n\|_2] \leqslant \mathcal{O}((1+\sigma)\sqrt{\ln(n)})$. Moreover, since:

$$\sum_{k=0}^{n-1} \gamma_k \alpha_k \mathbb{E}[\|\boldsymbol{\Lambda}_k\|_2] \leqslant \alpha \sum_{k=0}^{n-1} \gamma_k^2 \mathcal{O}(\sqrt{\ln(k)}) \leqslant \mathcal{O}(\sqrt{\ln(n)}) \sum_{k=0}^{n-1} \gamma_k^2,$$

it holds:

$$\mathbb{E}[\mathrm{CVio}_n^r] \leqslant \mathcal{O}(\ln^{3/2}(n)/\sqrt{n}).$$

This observation and (4.34) immediately give the following non-asymptotic result for the violation of the coupled constraint:

Theorem 4.8: *Suppose that*

$$\gamma_k = \Theta\left(\frac{1}{\sqrt{k}}\right),$$

then for every $r \in [R]$, we have the guarantee:

$$\mathbb{E}[\mathbf{g}_r(\overline{\boldsymbol{X}}_n)] \leqslant \mathcal{O}\left(\frac{\ln^{\frac{3}{2}}(n)}{\sqrt{n}}\right)$$

Remark 12: In the case where the mirror maps are different from the Euclidean projection, the bound in Theorem 4.8 might have weaker dimensional dependence. One reason is that the quantity $\Delta\psi$ (see (4.35)), hidden in the $\mathcal{O}$-notation in Theorem 4.8, might have in this case weaker dimension dependence. Another reason is that in this case, the constant σ might possesses similar behavior. To be more specific, consider the case where for all $i \in [N]$, $\mathcal{X}_i$ is a D_i-dimensional simplex and the case where the noise is a sequence of independent Gaussian normal vector in $\mathbb{R}^D$.

- In the case where for all $i \in [N]$, $\|\cdot\|_i$ is the 2-norm dual to itself, ψ_i is the Euclidean norm, and the mirror maps are Euclidean projections, it holds :

$$\Delta\psi = \sum_{i=1}^{N} \sqrt{D_i} \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{N} D_i.$$

- In comparison, we have in case for all $i \in [N]$, $\|\cdot\|_i$ is the 1-norm dual whose dual is the ∞-norm, ψ_i is the Gibbs entropy, and the mirror maps are logit choices, it holds:

$$\Delta\psi = \sum_{i=1}^{N} \ln(D_i) \quad \text{and} \quad \sigma^2 \leqslant \sum_{i=1}^{N} (2(\sqrt{\ln(D_i)} + \ln(D_i)) + 1),$$

  where $C > 0$ is an universal constant (for the latter see e.g. Example 2.7 in [116]).

4.6. Convergence Analysis for Ergodic Average

Theorem 4.8 guarantees decaying coupled constraints violation in expectation for the ergodic average of the population actions for step size γ_k and augmentation sequence α_k of order $\Theta(1/\sqrt{k})$. In contrast, Theorem 4.1 does not ensure[1] (see Remark 7) the a.s. fulfillment of the ergodic average in the asymptotic regime for the step size sequences of order $\Theta(1/\sqrt{k})$. To close this theoretical gap, we show in this section the convergence of the ergodic average for that class of step sizes. Our result in this direction is the following:

Theorem 4.9: *Suppose that* $\mathbf{v}$ *is strictly monotone , and suppose that there exists* $\sigma > 0$ *s.t.* $\sigma_k \leqslant \sigma$, *for all* $k \in \mathbb{N}$, *and that:*

$$\sum_{k=0}^{\infty} \gamma_k = \infty \quad \frac{\sum_{k=0}^{n-1} \gamma_k^2}{\sum_{k=0}^{n-1} \gamma_k} \to 0 \tag{4.36}$$

$$\frac{\sum_{k=0}^{n-1} \gamma_k \alpha_k}{\sum_{k=0}^{n-1} \gamma_k} \to 0 \tag{4.37}$$

$$\sum_{k=0}^{\infty} \frac{\gamma_k^2}{\left(\sum_{i=0}^{k} \gamma_i\right)^2} < \infty. \tag{4.38}$$

Then it holds $\lim_{n\to\infty} \overline{\boldsymbol{X}}_n = x_*$

Before we provide the proof of above Theorem, let us first give a discussion about step size and augmentation sequence fulfilling the requirement of the above Theorem:

Remark 13: Clearly, step size sequence of order $\gamma_k = \Theta(1/k^p)$, where $p \in (0, 1]$ fullfils (4.36) and (4.38). Choosing $\alpha_k = \alpha\gamma_k$ for an $\alpha > 0$, one sees that (4.37) is also fulfilled.

Lemma 4.10: *Suppose that:*

$$\sum_{k=0}^{\infty} \gamma_k = \infty \quad \sum_{k=0}^{\infty} \frac{\gamma_k^2 \sigma_k^2}{\left(\sum_{i=0}^{k} \gamma_i\right)^2} < \infty$$

$$\frac{\boldsymbol{S}_n(\boldsymbol{x})}{\sum_{k=0}^{n-1} \gamma_k} \to 0 \quad a.s.\ k \to \infty$$

Proof: By the Cauchy-Schwarz inequality, by the assumption that $\mathcal{X}$ is compact, and by similar argumentation as in Remark 8, it follows that a.s.:

$$\sum_{k=0}^{\infty} \frac{\gamma_k^2 \mathbb{E}_k \left[\left| \tilde{\boldsymbol{M}}_{k+1} \right|^2 \right]}{\left(\sum_{i=0}^{k} \gamma_i\right)^2} < \infty \quad \text{a.s.}$$

Since $(\boldsymbol{S}_n(\boldsymbol{x}))_n$ is a martingale (see Lemma 4.3), we have from Theorem 2.9 (specifically statement (3.13)) the desired statement. ■

[1] The convergence of iterate implies the convergence of the ergodic average of the iterates

Lemma 4.11: *Suppose that $\sigma_k \leqslant \sigma$, for all k, and that:*

$$\frac{\sum_{k=0}^{n-1} \gamma_k^2}{\sum_{k=0}^{n-1} \gamma_k} \to 0, \quad n \to \infty. \tag{4.39}$$

Then it holds:

$$\frac{\boldsymbol{R}_n}{\sum_{k=0}^{n-1} \gamma_k} \to 0, \quad n \to \infty, \ a.s.$$

Proof: $(\|\boldsymbol{M}_n\|_*^2)_n$ is a submartingale with $\sup_{k\in\mathbb{N}} \mathbb{E}[\|\boldsymbol{M}_n\|_*^2] < \infty$. So, it follows from Theorem 2.9 the a.s. convergence of $(\|\boldsymbol{M}_k\|_*^2)_n$ and thus a.s. the existence of $A > 0$ s.t. $\sup_{n\in\mathbb{N}} \|\boldsymbol{M}_n\|_*^2 \leqslant A$. Combining this with (4.39), the result follows. ∎

Proof (Proof of Theorem 4.9): It holds:

$$\langle \overline{\boldsymbol{Z}}_n - \boldsymbol{z}_*, \tilde{\mathbf{v}}(\boldsymbol{z}_*)\rangle = \frac{\sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{Z}_k - \boldsymbol{z}_*, \tilde{\mathbf{v}}(\boldsymbol{z}_*)\rangle}{\sum_{k=0}^{n-1} \gamma_k} \geqslant \frac{\sum_{k=0}^{n-1} \gamma_k \eta_k}{\sum_{k=0}^{n-1} \gamma_k},$$

where the inequality follows from the monotonicity of $\tilde{\mathbf{v}}$. Thus it follows from (4.24) and the fact that $\mathcal{V}_n \geqslant \tilde{\mathrm{F}}(\boldsymbol{Z}_0, z_*)$:

$$\langle \overline{\boldsymbol{Z}}_n - \boldsymbol{z}_*, \tilde{\mathbf{v}}(\boldsymbol{z}_*)\rangle \geqslant -\frac{\tilde{\mathrm{F}}(\boldsymbol{Z}_0, \boldsymbol{z}_*)}{\tau_n} - \frac{\tilde{C}_1 \sum_{k=0}^{n-1} \gamma_k^2}{\tau_n} + \frac{\|\boldsymbol{\lambda}_*\|_2^2 \sum_{k=n_0}^{n-1} \gamma_k \alpha_k}{2\tau_n} + \frac{\boldsymbol{S}_n + \frac{2}{K}\boldsymbol{R}_n}{\tau_n}.$$

Lemma 4.10 and 4.11 give:

$$(\boldsymbol{S}_n + \frac{2}{K}\boldsymbol{R}_n)/\tau_n \to 0 \quad \text{a.s}$$

and consequently:

$$\liminf_{n\to\infty} \langle \overline{\boldsymbol{Z}}_n - \boldsymbol{z}_*, \tilde{\mathbf{v}}(\boldsymbol{z}_*)\rangle \geqslant 0.$$

Combining this statement with the fact that a.s.:

$$\limsup_{n\to\infty} \langle \overline{\boldsymbol{Z}}_n - \boldsymbol{z}_*, \tilde{\mathbf{v}}(\boldsymbol{z}_*)\rangle \leqslant 0,$$

which is implied by the fact that $\boldsymbol{z}_* \in \mathrm{SOL}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$, it follows that:

$$\lim_{n\to\infty} \langle \overline{\boldsymbol{Z}}_n - \boldsymbol{z}_*, \tilde{\mathbf{v}}(\boldsymbol{z}_*)\rangle = 0.$$

Since $\tilde{\mathbf{v}}$ is strictly monotone, it follows that a.s. $\lim_{n\to\infty} \overline{\boldsymbol{Z}}_n = z_*$ as desired. ∎

Remark 14 (Differences between Theorems 4.9 and 4.1): In contrast to Theorem 4.1, the proof of Theorem 4.9 does not rely on the reciprocity condition for the Fenchel coupling (Assumption 4.3). Furthermore, the requirement of Theorem 4.1 cannot be changed to the requirement of Theorem 4.9 so that a.s. compliance of the coupled resource constraints in the asymptotic limit also holds for step size sequences of order

$\gamma_k = \Theta(1/k^p)$ for all $p \in (0,1]$. The main reason is the Robbins-Siegmund argumentation in the proof of Theorem 4.1 (see Lemma 4.5)

4.7. Non-Asymptotic Bounds for the Distance to Variational Nash Equilibrium

The aim of this section is to investigate the convergence speed of noisy MDAL. In order to keep the argumentation simple, we assume again that $\mathbf{v}$ is strictly monotone, and denote the unique solution of $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^M_{\geqslant 0}, \tilde{\mathbf{v}})$ by z_*. Moreover, we assume throughout that (4.8) holds for all $n \geqslant 0$. A sufficient condition for (4.8) is given in the following

Lemma 4.12: *If we set (γ_n) non-decreasing and $\alpha_n = \alpha\gamma_n$ with:*

$$\gamma_0 \leqslant \frac{\sqrt{K}}{\sqrt{32}C_1} \quad \text{and} \quad \alpha \leqslant \frac{128C_1}{24K} \tag{4.40}$$

then (4.8) *holds for all $n \geqslant 0$.*

Proof: Let (γ_n) be a given non-increasing sequence in $\mathbb{R}_{\geqslant 0}$ and $\alpha_n = \alpha\gamma_n$, where $\alpha > 0$. Similar computation as in Remark 7 gives that (4.8) holds for all $n \geqslant 0$ if and only if:

$$2\alpha^2\gamma_n^2 + \frac{C_1^2}{K} - \frac{\alpha}{2} \leqslant 0, \quad \forall n \geqslant 0. \tag{4.41}$$

Notice that in order (4.41) holds, it is necessary that:

$$\gamma_n \leqslant \frac{\sqrt{K}}{\sqrt{32}C_1}, \quad \forall n \geqslant 0. \tag{4.42}$$

By the monotonicity of the step size sequence this holds if and only if this inequality holds for $n = 0$. Now, suppose that (4.42) holds. We have that (4.41) holds if and only if:

$$\left[\frac{\frac{1}{4} - \sqrt{\frac{1}{4} - 8\gamma_n^2\frac{C_1^2}{K}}}{4\gamma_n^2}\right]_+ \leqslant \alpha \leqslant \frac{\frac{1}{4} + \sqrt{\frac{1}{4} - 8\gamma_n^2\frac{C_1^2}{K}}}{4\gamma_n^2}. \tag{4.43}$$

Now, suppose that it holds:

$$\gamma_n \leqslant \frac{\sqrt{3K}}{C_1\sqrt{128}}, \quad \forall n \geqslant 0. \tag{4.44}$$

Again by monotonicity of the step size sequence, it is sufficient to require that (4.44) holds for $n = 0$. We have that (4.43) holds if and only if:

$$\alpha \leqslant \frac{\frac{1}{4} + \sqrt{\frac{1}{4} - 8\gamma_n^2\frac{C_1^2}{K}}}{4\gamma_n^2}, \quad \forall n \geqslant 0. \tag{4.45}$$

Since by (4.44) it follows that:

$$\frac{1}{4}+\sqrt{\frac{1}{4}-8\gamma_n^2\frac{C_1^2}{K}}\geqslant\frac{1}{2},$$

we have by the monotonicity of the stepsize sequence, that the following condition is sufficient for (4.45):

$$\alpha\leqslant\frac{1}{8\gamma_0^2}. \tag{4.46}$$

Clearly by (4.44) it is sufficient to assume the following in order that (4.46) holds:

$$\alpha\leqslant\frac{128C_1}{24K}. \tag{4.47}$$

■

4.7.1. Primal-Dual Gap

We define the so-called *primal-dual equilibrium gap* as follows:

$$\Xi(\boldsymbol{z}):=\langle\boldsymbol{z}_*-\boldsymbol{z},\tilde{\mathbf{v}}(\boldsymbol{z})\rangle.$$

The primal-dual equilibrium gap is suitable to measure the distance between a point in the primal-dual space and the variationally Nash equilibrium. Indeed by strict monotonicity of $\tilde{\mathbf{v}}$ and the fact that $\boldsymbol{z}_*\in\mathrm{VI}(\mathcal{X}\times\mathbb{R}^M_{\geqslant 0},\tilde{\mathbf{v}})$. it holds:

$$\Xi(\boldsymbol{z})>0,\quad\text{whenever}\quad(\boldsymbol{x},\boldsymbol{\lambda})=\boldsymbol{z}\neq\boldsymbol{z}_*.$$

Moreover in case that $\mathbf{v}$ is in addition $c_{\mathbf{v}}$-strongly monotone it follows by Proposition 4.15 that Ξ provides an estimate of the distance between the primal iterate of noisy MDAL and variationally Nash equilibrium:

$$\Xi(\boldsymbol{z})\geqslant\frac{c_{\mathbf{v}}}{2}\|\boldsymbol{x}-\boldsymbol{x}_*\|^2. \tag{4.48}$$

Instead giving an estimate for the primal-dual equilibrium gap of the iterate of MDAL, we give an estimate for the ergodic averages of the primal-dual equilibrium gap of noisy MDAL:

$$\overline{\Xi}_n^\gamma:=\frac{\sum_{k=0}^{n-1}\gamma_k\Xi(\boldsymbol{Z}_k)}{\sum_{k=0}^{n-1}\gamma_k},$$

If $\mathbf{v}$ is in addition $c_{\mathbf{v}}$-strongly monotone we can relate the ergodic averages with the distance between the primal iterate of MDAL and the unique variationally Nash equlibrium of Γ. Indeed, by (4.48) and convexity of the squared norm it yields:

$$\overline{\Xi}_n^\gamma\geqslant\frac{c_{\mathbf{v}}}{2}\frac{\sum_{k=0}^{n-1}\gamma_k\|\boldsymbol{X}_k-\boldsymbol{x}_*\|^2}{\sum_{k=0}^{n-1}\gamma_k}\geqslant\frac{c_{\mathbf{v}}}{2}\min\left\{\|\overline{\boldsymbol{X}}_n^\gamma-\boldsymbol{x}_*\|^2,\min_{k\leqslant n-1}\|\boldsymbol{X}_k-\boldsymbol{x}_*\|^2\right\},$$

where:

$$\overline{\boldsymbol{X}}_n^{\gamma} := \frac{\sum_{k=0}^{n-1} \gamma_k \boldsymbol{X}_k}{\sum_{k=0}^{n-1} \gamma_k}$$

denotes the ergodic average of the primal iterate of noisy MDAL.

4.7.2. Expectation and High Probability Bound

By Combining Theorem 4.2, the discussions made in (4.5) and Lemma 4.3, we immediately obtain a bound for the expectation of $\overline{\Xi}_n^{\gamma}$:

Theorem 4.13: *Suppose that* (4.8) *is true for all* $n \geqslant 0$. *Then it holds:*

$$\mathbb{E}[\overline{\Xi}_k^{\gamma}] \leqslant \frac{\tilde{\mathrm{F}}(\boldsymbol{z}_*, \boldsymbol{Z}_0) + \tilde{C}_1 \sum_{k=0}^{n-1} \gamma_k^2 + \sum_{k=0}^{n-1} \gamma_k \alpha_k \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} + \frac{2}{K} \sum_{k=0}^{n-1} \gamma_k^2 \mathbb{E}[\|\xi_k\|_*^2]}{\sum_{k=0}^{n-1} \gamma_k}.$$

Proof: Since (4.8) is true for all $n \geqslant 0$, it follows that:

$$\overline{\Xi}_k^{\gamma} := \frac{\sum_{k=0}^{n-1} \gamma_k \Xi(\boldsymbol{Z}_k)}{\sum_{k=0}^{n-1} \gamma_k} \leqslant \frac{\tilde{F}(\boldsymbol{z}_*, \boldsymbol{Z}_0) + \tilde{C}_1 \sum_{k=0}^{n-1} \gamma_k^2 + \sum_{k=0}^{n-1} \gamma_k \alpha_k \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} + S_n + (2/K) R_n}{\sum_{k=0}^{n-1} \gamma_k}. \tag{4.49}$$

Since (S_n) is a mean zero martingale, we obtain the desired result by taking the expectation in above inequality. ∎

Remark 15: Suppose that the noise is persistent in the sense that there exists $G > 0$ s.t.

$$\mathbb{E}[\|\xi_k\|_*^2] \leqslant G, \quad \forall k \in \mathbb{N}.$$

From above theorem, it follows that in case that the time horizon n is fixed we can choose constant step size and constant augmentation sequence depending on n, s.t.:

$$\mathbb{E}[\overline{\Xi}_n^{\gamma}] \leqslant \mathcal{O}(G n^{-\frac{1}{2}}).$$

Moreover for variable step size and augmentation sequence of order: $\mathcal{O}(1/n)$, we can achieve the bound:

$$\mathbb{E}[\overline{\Xi}_n^{\gamma}] \leqslant \mathcal{O}(G^2 \ln(n) n^{-\frac{1}{2}}).$$

The detailed explanation is defered to Appendix 4.10.

Let us now assume in addition that the martingale noise has light tail in the following sense:

Assumption 4.5: *For each* n, $\|\xi_{n+1}\|_*$ *is subgaussian given* $\mathcal{F}_n$, *i.e. there exists* $\sigma_n > 0$ *s.t.:*

$$\mathbb{E}_{n-1}\left[\frac{\|\xi_{n+1}\|_*^2}{\sigma_n^2}\right] \leqslant 2 \quad a.s.$$

This condition is in particular equivalent to the condition that the tail probability of

$\|\xi_{n+1}\|_*$ given $\mathcal{F}_n$ is dominated (up to an absolute constant) by the tail probability of a normal random variable with variance σ_n^2. For detailed treatment of this aspect, we refer to [117].

With this additional structure of the noise, we can refine the bound given in Theorem 4.13 as follows:

Theorem 4.14: *With probability* $1-\delta$*, we have:*

$$\Xi_k^\gamma \leqslant \frac{\tilde{F}(\boldsymbol{z}_*, \boldsymbol{Z}_0) + \tilde{C}_1 \sum_{k=0}^{n-1} \gamma_k^2 + \sum_{k=0}^{n-1} \gamma_k \alpha_k \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} + (2/K) C_{\psi_1} \sum_{k=0}^{n-1} \gamma_k^2 \sigma_k^2}{\sum_{k=0}^{n-1} \gamma_k} \tag{4.50}$$

$$+ \frac{\sqrt{2\ln(2/\delta) C_{\psi_2} C_{\mathcal{X}}^2 \sum_{k=0}^{n-1} \gamma_k^2 \sigma_k^2} + \ln(2/\delta) C_{\psi_1} \sup_{k\in[n]} \sigma_{k-1}^2}{\sum_{k=0}^{n-1} \gamma_k}, \tag{4.51}$$

where:

$$C_{\psi_1} = 4e^{1+(1/e)} \quad \textit{and} \quad C_{\psi_2} = 2/e$$

The proof of this theorem is quite technical. Thus we defer it to Subsection 4.9.4. One can show that by Assumption 4.5, it follows that:

$$\mathbb{E}_{n-1}[\|\xi_{n+1}\|_*^2] \lesssim \sigma_n^2. \tag{4.52}$$

Thus the first part (4.50) of the upper bound given in Theorem 4.14 coincides (up to an absolute constant) with the expectation bound given in Theorem 4.13. So the additional price we pay for analyzing the probability instead of the expectation is given in the second part (4.51).

Remark 16: Suppose that the noise is persistent in the sense that:

$$\sigma_k = \sigma, \quad \forall k.$$

From above Theorem, it follows that in case that the time horizon n is fixed choosing constant step size and constant augmentation sequence depending on n yields with probability $1-\delta$

$$\Xi_k^\gamma \leqslant \mathcal{O}\left(\sqrt{\ln(2/\delta)\sigma^2/n}\right).$$

Moreover for variable step size and augmentation sequence of order $\mathcal{O}(1/n)$, with probability $1-\delta$ we can achieve the bound:

$$\Xi_n^\gamma \leqslant \mathcal{O}\left(\ln(2/\delta)\sigma^2(\ln(n)/\sqrt{n})\right).$$

The detailed explanation is defered to Appendix 4.10

4.8. Numerical Experiment

In this section, we numerically investigate the behavior of Algorithm 3 in case the underlying game model is quadratic. Quadratic game is a popular model in several applications (see e.g. [76, 84, 102]) such as competitive markets, cognitive radio network, charging of electric vehicles, and congestion control of road network. Our focus lies on the difference of the sustainable resource behavior, between the population's states and the their ergodic averages, and between different choices of mirror map.

4.8.1. Setting

Exponential Weights Online Learning in Quadratic Game We consider N agents whose task is to allocate a certain amount of tasks to R resources. The strategy space of agent i corresponds to the simplex $\varDelta := \left\{\boldsymbol{x}^{(i)} \in \mathbb{R}^R_{\geqslant 0} : \sum_{r=1}^R \boldsymbol{x}_k^{(i)} = 1\right\}$. For a strategy $\boldsymbol{x}^{(i)} \in \varDelta$, $x_r^{(i)}$ stands for the proportion of tasks agent i assigns to resource $r \in [R]$. Moreover, we have, in this case, $D = R$. The cost function of player i is quadratic and given by

$$J^{(i)}(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(-i)}) = \frac{1}{2}\langle \boldsymbol{x}^{(i)}, \mathbf{Q}\boldsymbol{x}^{(i)}\rangle + \langle \mathbf{C}\sigma(\boldsymbol{x}) + c^i, \boldsymbol{x}^{(i)}\rangle,$$

where $\sigma(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^N \boldsymbol{x}^{(i)}$ where $c_i \in \mathbb{R}^D$, $\mathbf{Q} \in \mathbb{R}^{D\times D}$ and $\mathbf{C} \in \mathbb{R}^{D\times D}$ are positive semi-definite, and either $\mathbf{Q}$ or $\mathbf{C}$ are positive definite. In order to apply our method, we set $\mathrm{u}^{(i)}(\boldsymbol{x}) = -J^{(i)}(\boldsymbol{x})$. The corresponding gradient mapping is given by:

$$\mathbf{v}(\boldsymbol{x}) = -\left[(\mathbf{I}_N \otimes \mathbf{Q} + \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^{\mathrm{T}} \otimes C)\boldsymbol{x} + c + \frac{1}{N}(\mathbf{I}_N \otimes C^{\mathrm{T}})\boldsymbol{x}\right]$$

where $\otimes$ denotes the Kronecker product between two matrices, and where $\mathbf{1}_N$ denotes a vector in $\mathbb{R}^N$ whose entries are one and $\mathbf{I}_N$ denotes the identity matrix in $\mathbb{R}^{N\times N}$.

Game Parameter As mirror map of the agents, we use either the Euclidean projection onto the simplex or the logit choice. In order to avoid numerical overflow in the implementation of the logit choice we use the log-sum trick. In the simulation, we set:

$$\mathbf{Q} = 2\sqrt{\tilde{\mathbf{Q}}^{\mathrm{T}}\tilde{\mathbf{Q}}} + \mathbf{I}_D,$$

where the entries of $\tilde{\mathbf{Q}}$ is chosen independently normal distributed. Moreover we consider the case where :

$$\mathbf{C} = 4\mathbf{I}_D \quad \text{and} \quad c = 0.$$

We adapt as the model for resource constraints the affine constraints as described in Remark 5 with:

$$\mathbf{A} = 4\mathbf{E}_D \quad \text{and} \quad b = d\mathbf{1}_D,$$

for different $d > 0$, where $\mathbf{E}_D$ is the matrix having entries $= 1$ As the model for the stochastic feedback, we use:

$$\text{the Gaussian vector with the covariance matrix } \sigma^2 \mathbf{I}_D,$$

where either $\sigma = 0$ or $\sigma = 5$. Throughout the simulations, we choose the step size sequence :

$$\gamma_n = \frac{0.5}{\sqrt{n+1}}$$

and the augmentation sequence:

$$\alpha_n = \alpha \gamma_n \quad \text{with} \quad \alpha = 5.$$

We plot the corresponding resource average negative clipped constraints violation (RANCCV) both of the ergodic average of the population's dynamic, i.e., $\sum_{r=1}^{R} [\mathbf{g}_r(\overline{\boldsymbol{X}}_n)]_+ / R$, and of the population's dynamic, i.e., $\sum_{r=1}^{R} [\mathbf{g}_r(\boldsymbol{X}_n)]_+ / R$.

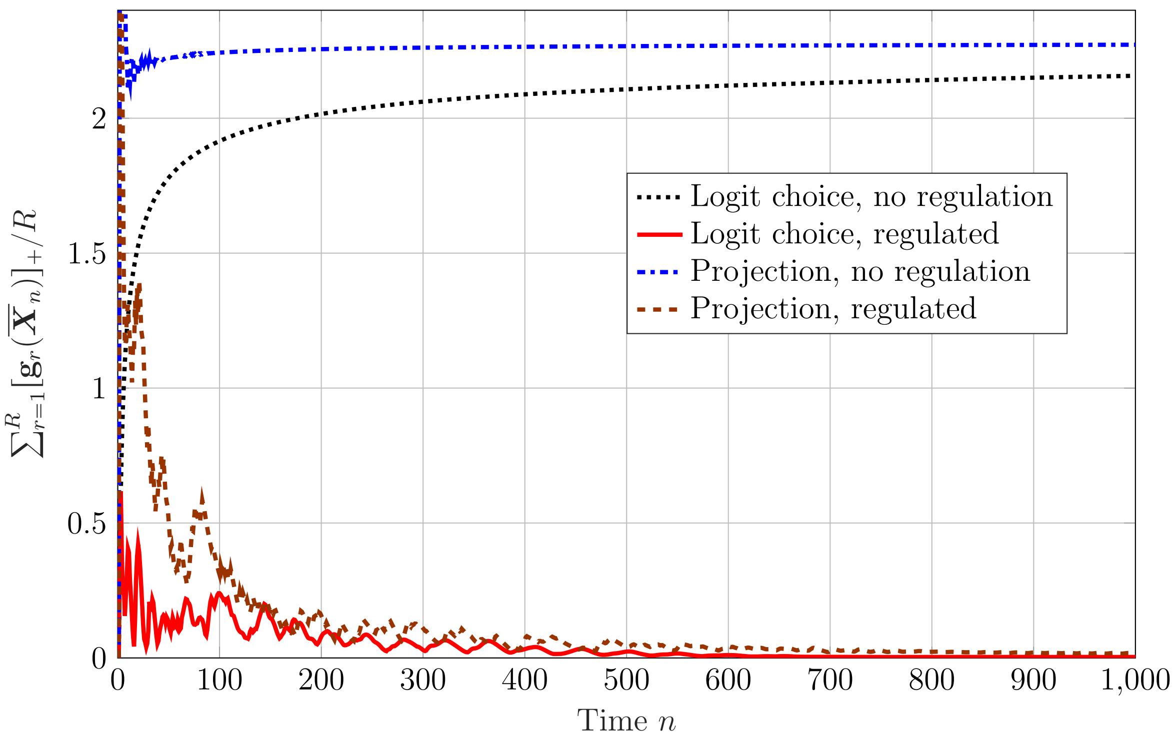


Figure 4.1.: RANCCV of the ergodic average of the population's dynamic for $D = R = 20$, $N = 50$, $d = 10.5$, and $\sigma = 0$.

4.8.2. Evaluation

Price regulation vs. Anarchy - RANCCV We first evaluate the simulation results for the case of no feedback noise, i.e., $\sigma = 0$ (Figure 4.1–4.5), in the case of no feedback noise, i.e., $\sigma = 0$. As apparent from Figure 4.1, control of selfish agents is crucial for sustainable

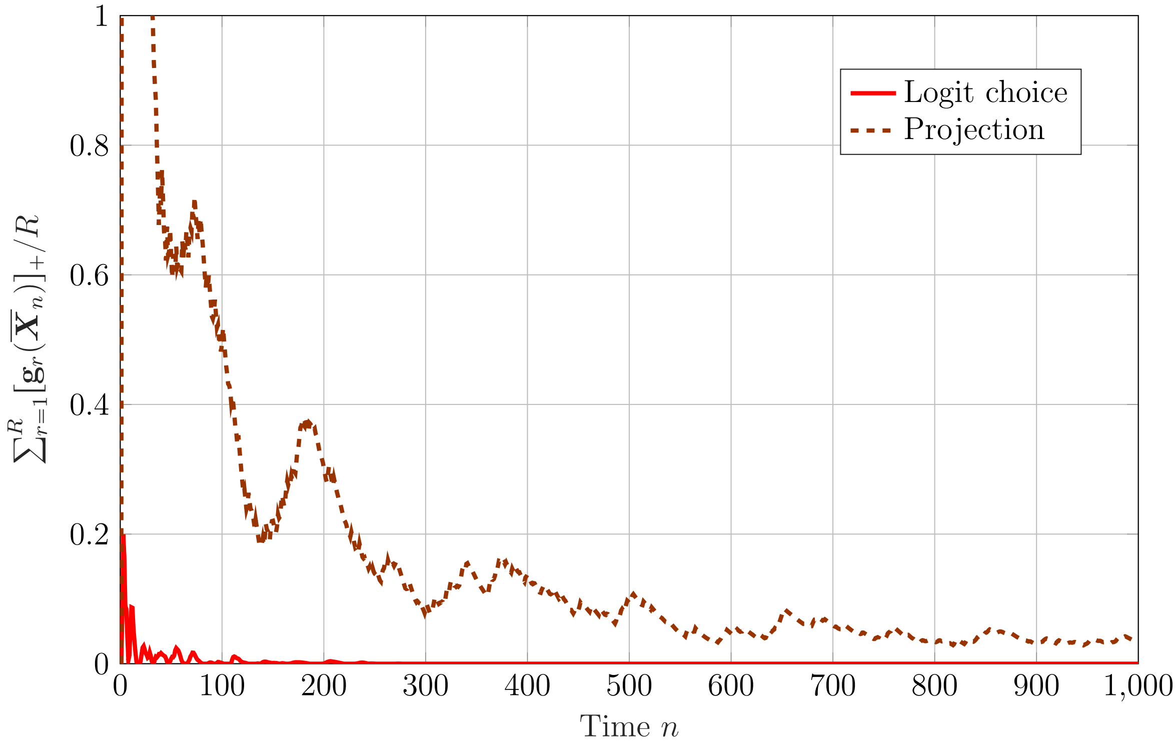


Figure 4.2.: RANCCV of the ergodic average of the population's dynamic for $D = R = 50$, $N = 100$, $d = 8.5$, and $\sigma = 0$.

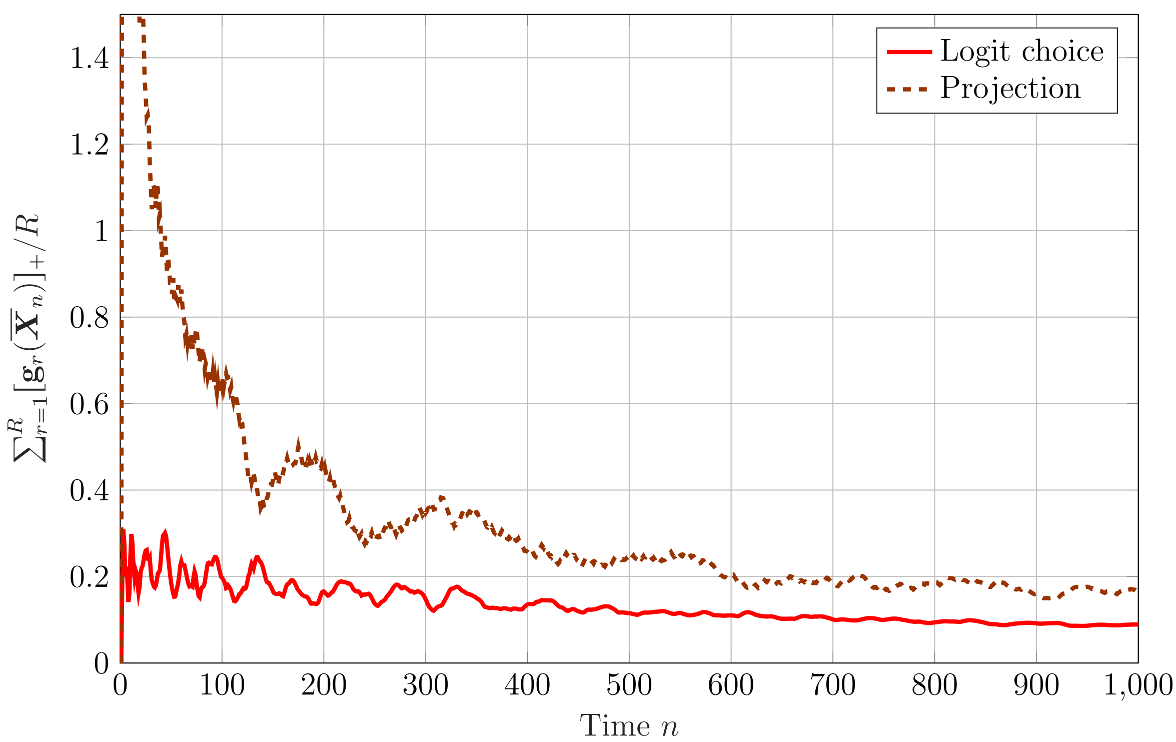


Figure 4.3.: RANCCV of the ergodic average of the population's dynamic for $D = R = 50$, $N = 100$, $d = 8$, and $\sigma = 0$

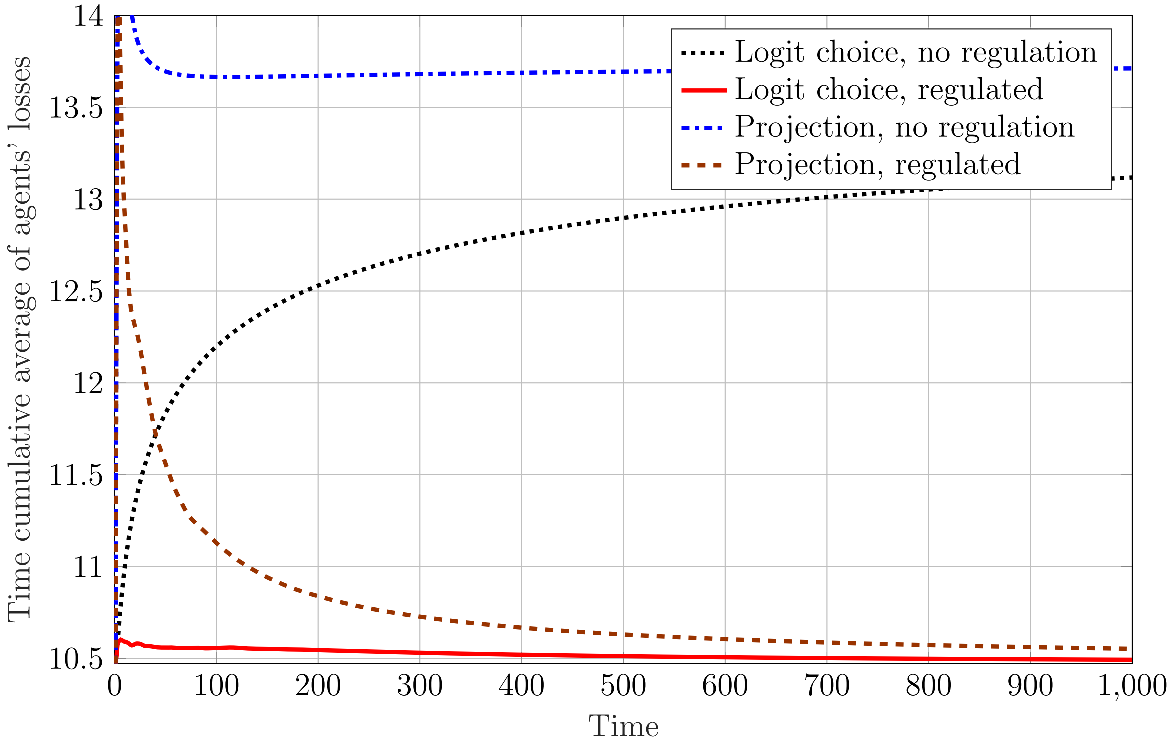


Figure 4.4.: Comparison of Time cumulative of average agents' utilities for mirror ascent dynamic with (MAARP) and without pricing regulation.

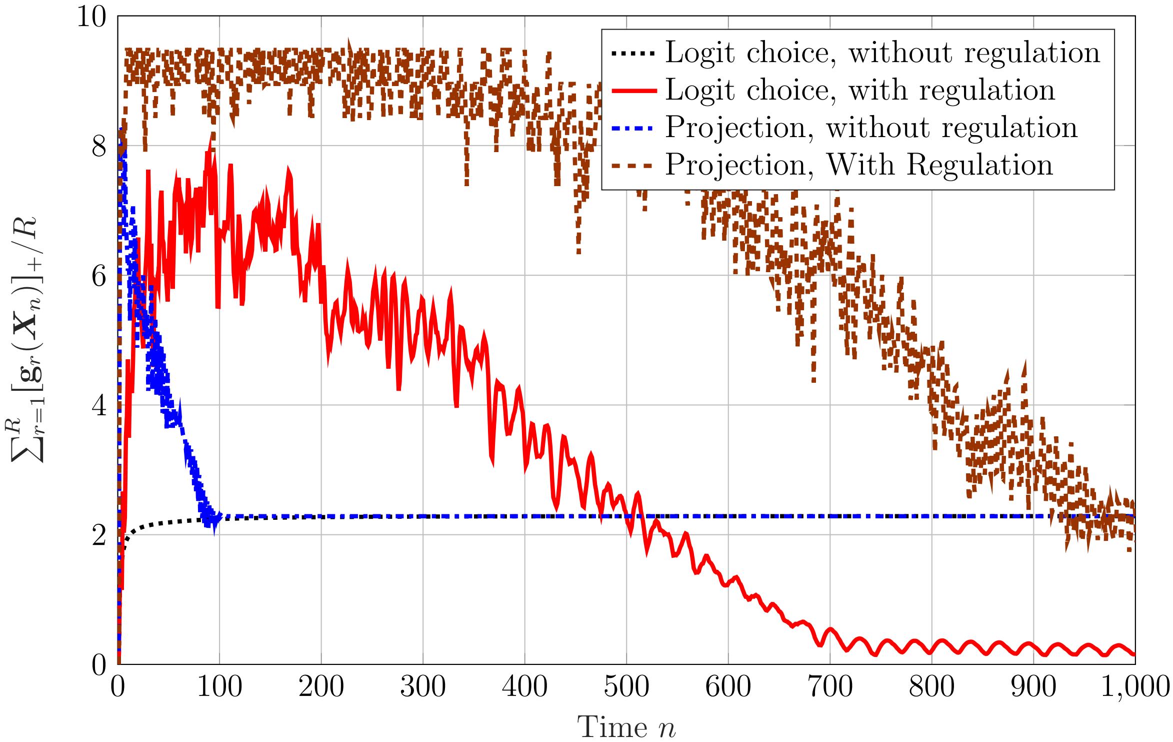


Figure 4.5.: RANCCV of the population's dynamic for $D = R = 20$, $N = 50$, $d = 10.5$, and $\sigma = 0$.

resource behaviour, since the average of the negative clipped resource constraints violation in the pure anarchistic case (black dotted - and blue dash-dotted line) (4.2) is significantly higher than the case where MAARP is applied (red - and brown dashed line). We observe that better results yields if we use the logit choice (red line) instead of the usual Euclidean projection (brown dashed line). This effect is more pronounced in a high strategy space dimension than in the low one, cf. Figure 4.1 and Figures 4.2–4.3, which is aligned with the discussion made in Remark 12.

Price regulation vs. Anarchy - Utilities Given the previous discussion, one may think that the reduction of constraint violation comes with a reduction of the population's welfare. For this reason, we investigate the average of agents' losses, i.e. $\sum_{i=1}^{N} J^{(i)}(\overline{\boldsymbol{X}}_n)/N$, numerically . We plot the time-average[2] of this quantity in Figure 4.4. There we observe that MAARP, in contrast to the pure anarchistic case, not only promotes sustainable behaviour but also reduces the average loss, and therefore improves the populations' welfare. Furthermore, we see that the use of a mirror map other than Euclidean projection also improves not only sustainable behavior but also the population's welfare.

Ergodic average vs. Indeed Trajectory Figure 4.2 depicts the behavior of the indeed population's dynamic. Therein we can observe the tendency of the decaying amount of violations of the resource constraints. However, the corresponding decay might be much slower in comparison to decay of the ergodically time-averaged population's dynamic (see Figure 4.1).

Noise-Robustness of MAARP with Logit choice Now we evaluate the simulation results for the case of feedback noise with $\sigma = 5$ (Figures 4.6 and 4.7) and 500 noisy samples (i.i.d). One can see that using logit choice effects in average better results than using Euclidean projection as forecasted by the discussion made in Remark 12. Moreover, one can observe that the RANCCVC is more volatile in the case where the Euclidean projection is used compared to the case where the mirror map is the logit choice. Comparing Figures 4.6 and 4.7, it is apparent that using the ergodic average yields significantly lower and less volatile amount of resource constraint violations.

Comparison to the State of the Art At last, we give a comparison of our method to some existing methods comparable to ours, i.e., the primal-dual method (see e.g. [55]) and the asymmetric projection (AP) (see e.g., Algorithm 2 [84]). The primal-dual method can be seen as the MAARP without augmentation (i.e., $\alpha = 0$). Moreover, we give the primal-dual method leverage by equipping it with the logit choice instead of the Euclidean projection. We plot the corresponding RANCCVC in Figure 4.8. There, one can see

[2]The reason that we plot the time-average instead of the quantity itself is to make the performance distinction between MAARP with logit choice and with Euclidean projection clearer.

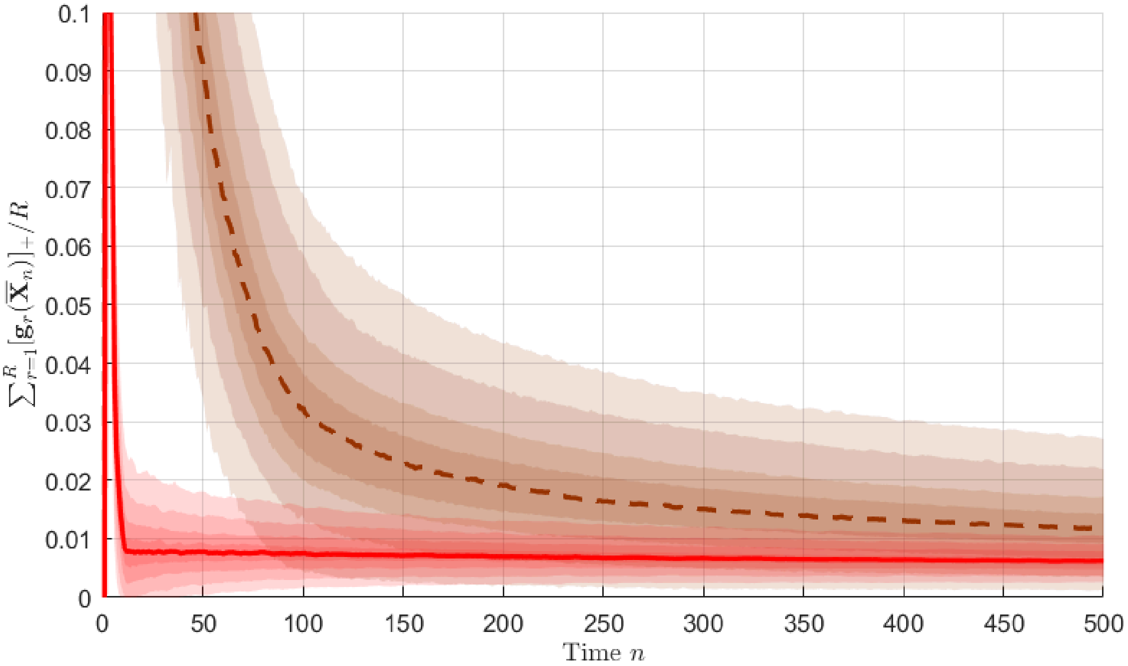


Figure 4.6.: RANCCV of the ergodic average of the population's dynamic for $D = R = 20$, $N = 50$, $d = 10.5$, $\sigma = 5$, and sample size $= 500$. Red line corresponds to the sample average of RANCCVC in the logit choice case and brown dashed line resp. in the Euclidean case. Shaded areas are each corresponds to 25%-, 50%-, 75%-, and 90%-percentile.

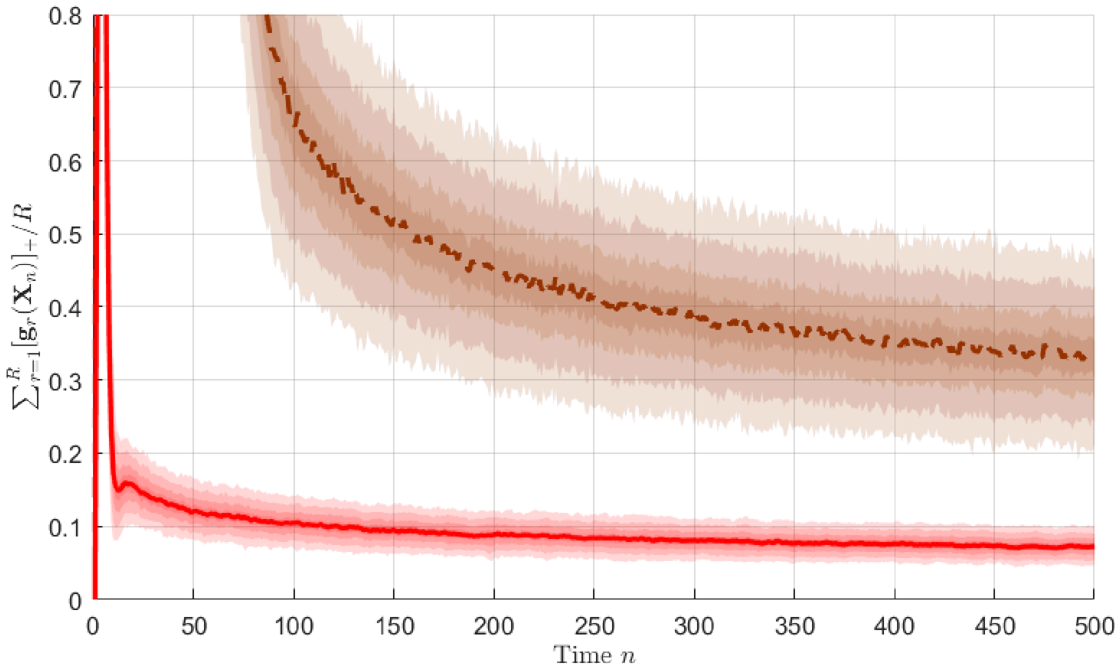


Figure 4.7.: RANCCV of the population's dynamic for $D = R = 50$, $N = 100$, $d = 10.5$, $\sigma = 5$, and sample size $= 500$. Red line corresponds to the sample average of RANCCVC in the logit choice case and brown dashed line resp. in the Euclidean case. Shaded areas are each corresponds to 25%-, 50%-, 75%-, and 90%-percentile.

that the MARRP with logit choice (red line) outperforms the primal-dual method (yellow dash-dotted), and the best result yields if one uses AP (purple dotted). However, the dual variable update of AP requires, in contrast to MAARP, two consecutive congestion states of the resources and can not be implemented parallelly with the population's strategy update (see Remark 6). Besides this fact, one can see from the plot of the average of agents' losses in Figure 4.9, that the excellent performance of RANCCVC of AP comes with the increase of agents' losses.

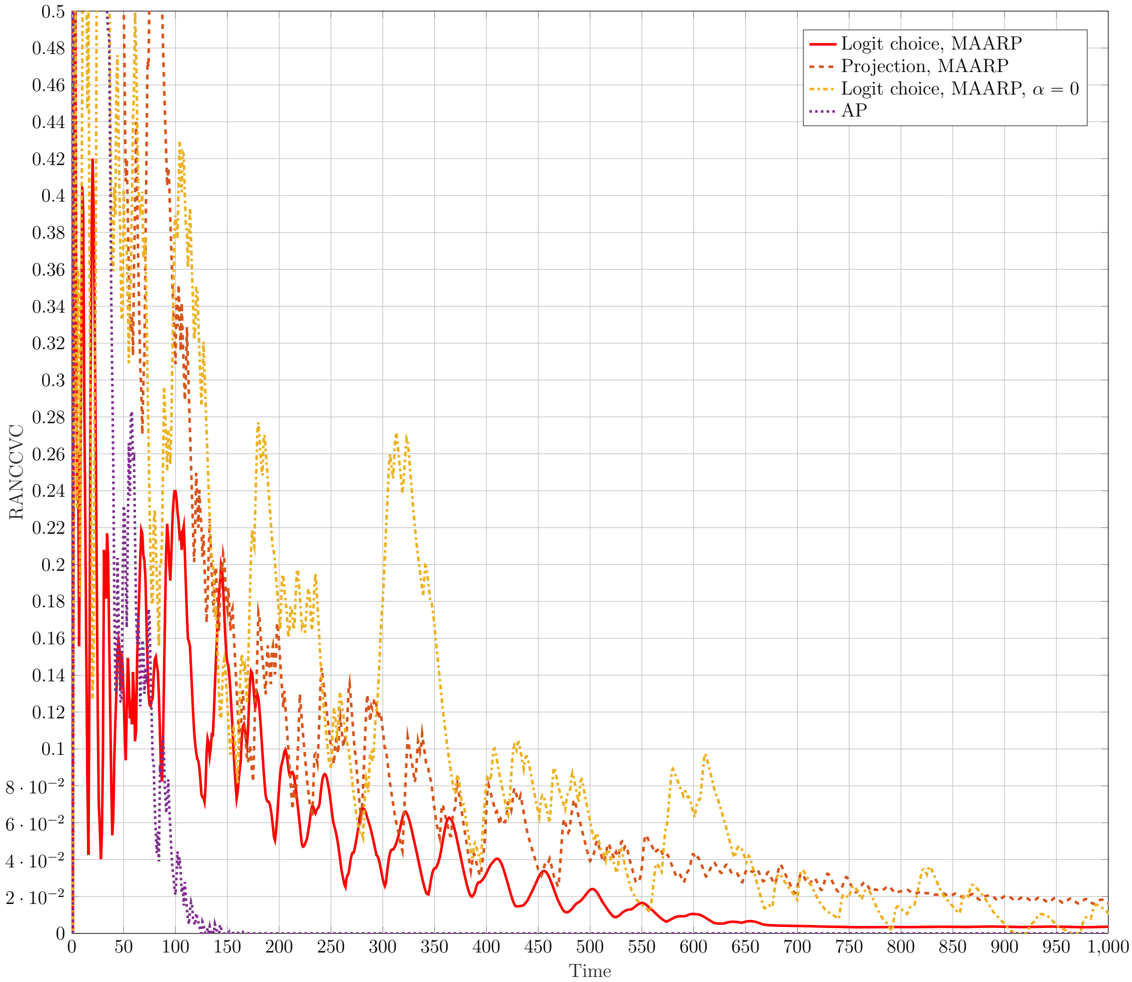


Figure 4.8.: Comparison of RANCCVC of MAARP with state of the art.

4.9. Appendix

4.9.1. Monotonicity of the KKT Operator

Proposition 4.15: *Suppose that* $\mathbf{g}_i$ *is convex for all* $i \in [N]$*. It holds:*

$$\langle \alpha_1 - \alpha_2, \tilde{\mathbf{v}}(\alpha_1) - \tilde{\mathbf{v}}(\alpha_2) \rangle \leqslant \langle \boldsymbol{x}_1 - \boldsymbol{x}_2, \mathbf{v}(\boldsymbol{x}_1) - \mathbf{v}(\boldsymbol{x}_2) \rangle, \tag{4.53}$$

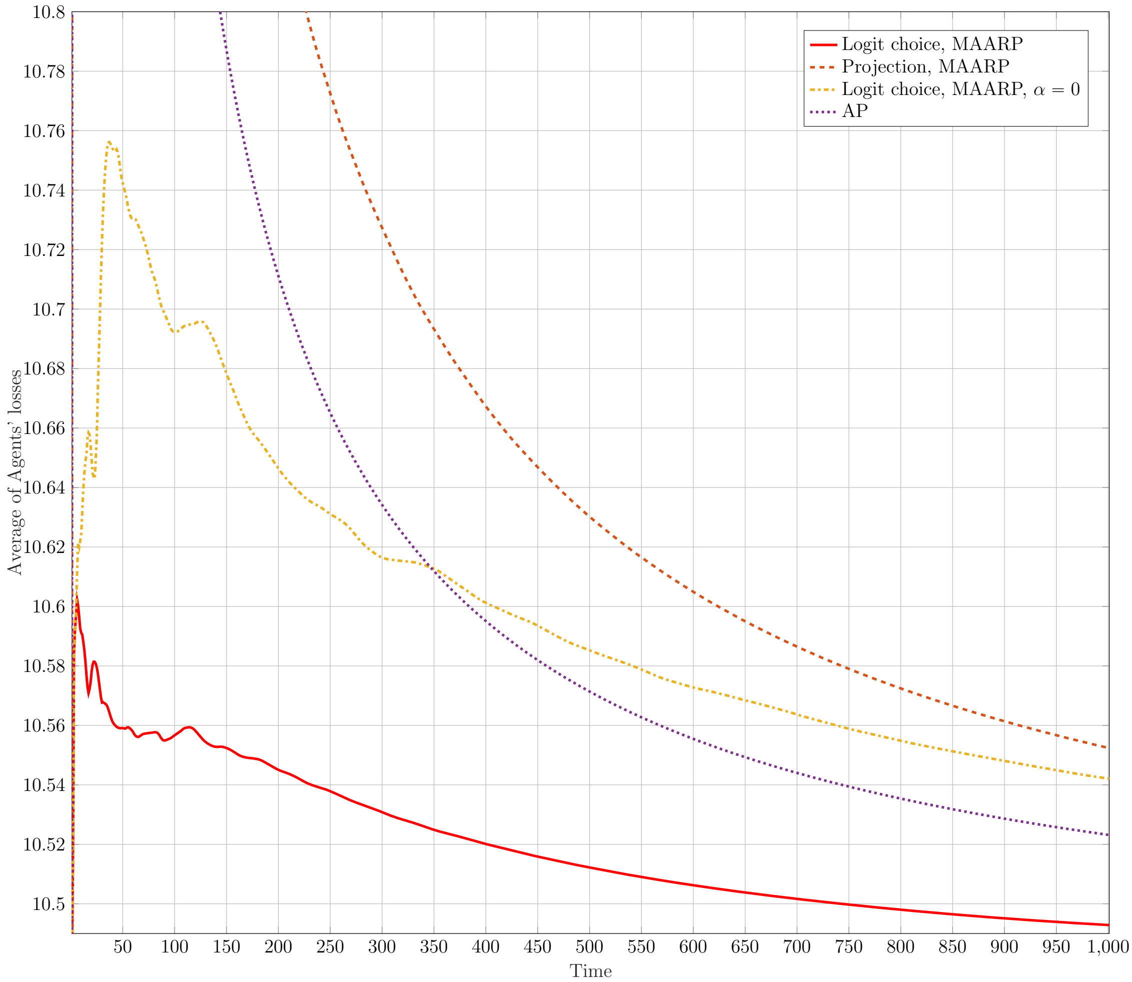


Figure 4.9.: Comparison of the average agents' losses of MAARP with state of the art.

for all $\alpha_i := (\boldsymbol{x}_i, \boldsymbol{\lambda}_i) \in \mathcal{X} \times \mathbb{R}^M_{\geqslant 0}$, $i = 1, 2$. *Consequently if* $\mathbf{v}$ *is (resp. strictly) monotone then* $\mathbf{v}$ *is (resp. strictly) monotone. Moreover if* $\mathbf{v}$ *is c-strongly monotone, then :*

$$\langle \alpha_1 - \alpha_2, \tilde{\mathbf{v}}(\alpha_1) - \tilde{\mathbf{v}}(\alpha_2)\rangle \leqslant -\frac{c}{2}\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2,$$

Proof: Straightforward computations yields:

$$\begin{aligned}&\langle \alpha_1 - \alpha_2, \tilde{\mathbf{v}}(\alpha_1) - \tilde{\mathbf{v}}(\alpha_2)\rangle = \langle \boldsymbol{x}_1 - \boldsymbol{x}_2, \mathbf{v}(\boldsymbol{x}_1) - \mathbf{v}(\boldsymbol{x}_2)\rangle \\ &- \langle \boldsymbol{x}_1 - \boldsymbol{x}_2, [\nabla \mathbf{g}(\boldsymbol{x}_1)]^T \boldsymbol{\lambda}_1 - [\nabla \mathbf{g}(\boldsymbol{x}_2)]^T \boldsymbol{\lambda}_2\rangle + \langle \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2, \mathbf{g}(\boldsymbol{x}_1) - \mathbf{g}(\boldsymbol{x}_2)\rangle.\end{aligned}$$

Since $\mathbf{g}_i$ is convex for all $i \in [M]$, it holds:

$$\begin{aligned}\langle \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2, \mathbf{g}(\boldsymbol{x}_1) - \mathbf{g}(\boldsymbol{x}_2)\rangle &= \sum_i \boldsymbol{\lambda}_1^{(i)}(\mathbf{g}_i(\boldsymbol{x}_1) - \mathbf{g}_i(\boldsymbol{x}_2)) - \sum_i \boldsymbol{\lambda}_2^{(i)}(\mathbf{g}_i(\boldsymbol{x}_1) - \mathbf{g}_i(\boldsymbol{x}_2)) \\ &\leqslant \sum_i \boldsymbol{\lambda}_1^{(i)}\langle \boldsymbol{x}_1 - \boldsymbol{x}_2, \nabla \mathbf{g}_i(\boldsymbol{x}_1)\rangle - \sum_i \boldsymbol{\lambda}_2^{(i)}\langle \boldsymbol{x}_1 - \boldsymbol{x}_2, \nabla g_i(\boldsymbol{x}_2)\rangle \\ &= \langle \boldsymbol{x}_1 - \boldsymbol{x}_2, [\nabla \mathbf{g}(\boldsymbol{x}_1)]^T \boldsymbol{\lambda}_1 - [\nabla \mathbf{g}(\boldsymbol{x}_2)]^T \boldsymbol{\lambda}_2\rangle\end{aligned}$$

Combining all both computations, we obtain (4.53). Now, if $\tilde{\mathbf{v}}$ is (resp. strictly) monotone, then we have $\langle \boldsymbol{x}_1 - \boldsymbol{x}_2, \mathbf{v}(\boldsymbol{x}_1) - \mathbf{v}(\boldsymbol{x}_2)\rangle \leqslant 0$ (resp, < 0). this observation and (4.53) gives the last statement. ∎

4.9.2. Nash Equilibrium & Variational Inequality

Decoupling the Coupled Constraints via of Lagrangian

In order to investigate $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$, it is convenient to extend $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ to $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_+, \tilde{\mathbf{v}})$, where $\tilde{\mathbf{v}} : \mathcal{X} \times \mathbb{R}^R_+$,

$$\tilde{\mathbf{v}} : \mathcal{X} \times \mathbb{R}^R_+ \to \mathbb{R}^{D+R}, (\boldsymbol{x}, \boldsymbol{\lambda}) \mapsto \left[\mathbf{v}(\boldsymbol{x}) - [\nabla \mathbf{g}(\boldsymbol{x})]^{\mathrm{T}} \boldsymbol{\lambda}, \mathbf{g}(\boldsymbol{x})\right]^{\mathrm{T}}. \tag{4.54}$$

The advantage of this method is the decoupling of the constraint set. Specifically, employing this procedure, we only have to work with the constraint set $\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}$ with product structure rather than with $\mathcal{Q}$.

The following Proposition is toward that direction:

Proposition 4.16: *Suppose that the Assumptions 3.3, 4.1 and 4.2 holds. Then the following statements are equivalent:*

1. $\overline{\boldsymbol{x}} \in \mathcal{Q}$ *is a solution of* $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$
2. *There exists* $\overline{\boldsymbol{\lambda}} \in \mathbb{R}^R_{\geqslant 0}$ *s.t.* $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ *is a solution of* $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_+, \tilde{\mathbf{v}})$.

Proposition 4.16 tells us that in order to find a variational Nash equilibrium of Γ (in case it exists), it is sufficient to find the solution of $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_+, \tilde{\mathbf{v}})$. Moreover, if $\mathbf{v}$ is additionaly strictly monotone, we can strengthen the statement given in that proposition as follows:

Proposition 4.17: *Suppose that Assumptions 3.3, 4.1 and 4.2 hold. If* $\mathbf{v}$ *is strictly monotone, then there exists a unique solution* $\overline{\boldsymbol{x}}$ *of* $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ *and* $\overline{\boldsymbol{\lambda}} \in \mathbb{R}^R_{\geqslant 0}$ *s.t.* $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ *is a unique solution of* $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$.

Proof: Proposition 4.16 ensures the existence of such a pair $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$. Moreover, since $\mathbf{v}$ is strictly monotone, it follows from Proposition 4.15 that $\tilde{\mathbf{v}}$ is strictly monotone. Consequently Proposition 2.6 ensures the uniqueness of $\mathrm{VI}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}})$ ■

The idea of relating constrained convex optimization problem to unconstrained convex optimization problem applies also to variational inequality. For this sake, we define the notion of KKT system relative to the variational inequality $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ by:

$$\begin{aligned} &\mathbf{v}(\boldsymbol{x}) - \boldsymbol{\lambda}^T \nabla \mathbf{g}(\boldsymbol{x}) = 0 \\ &0 \leqslant \boldsymbol{\lambda} \perp \mathbf{g}(\boldsymbol{x}) \leqslant 0 \\ &\boldsymbol{x} \in \mathcal{X}. \end{aligned} \tag{4.55}$$

This system is related to the KKT system:

$$\begin{aligned} &\nabla f(\boldsymbol{x}) - \nabla g(\boldsymbol{x})^T \boldsymbol{\lambda} = 0 \\ &0 \leqslant \boldsymbol{\lambda} \perp g(\boldsymbol{x}) \leqslant 0 \\ &x \in \mathcal{D} \end{aligned} \tag{4.56}$$

which corresponds to the constrained optimization problem:

$$\max_{\boldsymbol{x} \in \mathcal{D}} f(\boldsymbol{x}) \quad \text{s.t. } g(\boldsymbol{x}) \leqslant 0, \tag{4.57}$$

where $f : \mathcal{D} \to \mathbb{R}$, and $g : \mathcal{D} \to \mathbb{R}^M$ continuously differentiable and $\mathcal{D}$ is a non-empty subset of $\mathbb{R}^D$.

Proposition 4.18: *Suppose that Slater's condition holds. Then* $\overline{\boldsymbol{x}}$ *solves* $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ *if and only if there exists* $\overline{\boldsymbol{\lambda}}$ *s.t.* $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ *solves the KKT system* (4.55)

Proof: $\overline{\boldsymbol{x}}$ solves $\mathrm{VI}(\mathcal{Q}, \mathbf{v})$ if and only if:

$$\langle \boldsymbol{x}, \mathbf{v}(\overline{\boldsymbol{x}}) \rangle \leqslant \langle \overline{\boldsymbol{x}}, \mathbf{v}(\overline{\boldsymbol{x}}) \rangle, \quad \forall \boldsymbol{x} \in \mathcal{Z}.$$

In turn, the latter is equivalent with:

$$\overline{\boldsymbol{x}} \in \arg\max_{\boldsymbol{x} \in \mathcal{Q}} \langle \boldsymbol{x}, \mathbf{v}(\overline{\boldsymbol{x}}) \rangle.$$

The KKT system corresponds to above program is given by (4.55) (cf. (4.57)). Since $\boldsymbol{x} \mapsto \langle \boldsymbol{x}, F(\overline{\boldsymbol{x}}) \rangle$ is trivially convex on $\mathcal{X}$, and the Slater's constraint qualification is fulfilled, then the desired statement follows since for constrained convex optimization problem the set of the solution of the KKT system coincides with the set of solutions of the optimization problem. ■

In order to obtain Proposition 4.16 all we need is to get rid of the condition $\boldsymbol{\lambda} \perp \mathbf{g}(\boldsymbol{x})$ in the KKT system for VI (see (4.55)). This is done in the following:

Proof (Proof of Proposition 4.16): Since there are no explicit inequality constraint in $\mathcal{X} \times \mathbb{R}^M_{\geqslant 0}$ then $\mathcal{X} \times \mathbb{R}^M_{\geqslant 0}$ fulfills the slater's CQ. It follows from Proposition 4.18 that $\mathrm{SOL}(\mathcal{X} \times \mathbb{R}^M_{\geqslant 0}, \tilde{\mathbf{v}})$ coincides with the solution of the KKT system:

$$
\begin{aligned}
&\mathbf{v}(\boldsymbol{x}) - \nabla \mathbf{g}(\boldsymbol{x})\boldsymbol{\lambda} = 0 \\
&\mathbf{g}(\boldsymbol{x}) + \mu = 0 && (4.58)\\
&0 \leqslant \mu \perp -\boldsymbol{\lambda} \leqslant 0 && (4.59)\\
&\boldsymbol{x} \in \mathcal{X} \quad \boldsymbol{\lambda} \in \mathbb{R}^M_{\geqslant 0}. && (4.60)
\end{aligned}
$$

Setting (4.58) into (4.59) we obtain that $\mathrm{SOL}(\mathcal{X} \times \mathbb{R}^M_{\geqslant 0}, \tilde{\mathbf{v}})$ coincides with the solution of KKT system:

$$
\begin{aligned}
&\mathbf{v}(\boldsymbol{x}) - \nabla \mathbf{g}(\boldsymbol{x})\boldsymbol{\lambda} = 0 \\
&0 \leqslant \boldsymbol{\lambda} \perp \mathbf{g}(\boldsymbol{x}) \leqslant 0 \\
&\boldsymbol{x} \in \mathcal{X}. && (4.61)
\end{aligned}
$$

For the final step, notice that the solution of this KKT system is equivalent to $\mathrm{SOL}(\mathcal{Q}, \mathbf{v})$ since the Slater's CQ for $\mathcal{Q}$ is fulfilled. ■

4.9.3. Non-explosiveness of the Iterate of MDAL

Since $\boldsymbol{X}_n \in \mathcal{X}$ and $\mathcal{X}$ is compact, it follows immediately that $\|\boldsymbol{X}_n\| < \infty$. In order to proof the non-explosiveness of the dual iterate of MAARP, we need the following known result:

Proposition 4.19 (Discrete Gronwall's Inequality): *Let (y_n), (f_n), and (g_n) be non-negative sequences. If for $n \geqslant 0$ it holds:*

$$y_n \leqslant f_n + \sum_{k<n} g_k y_k,$$

then:

$$y_n \leqslant f_n + \sum_{k<n} g_k f_k \exp\Big(\sum_{k<j<n} g_j\Big)$$

In order to show the non-explosiveness of $\boldsymbol{\Lambda}_n$, notice first that from (4.20), it follows that:

$$\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2 \leqslant \|\boldsymbol{\Lambda}_0 - \boldsymbol{\lambda}\|_2^2 + 2\sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{g}(\boldsymbol{X}_k)\rangle + \sum_{k=0}^{n-1} \gamma_k \alpha_k \|\boldsymbol{\lambda}\|_2^2 \\ + 4\sum_{k=0}^{n-1} \gamma_k^2 (C_3^2 + \alpha_k^2 \|\boldsymbol{\Lambda}_k\|_2^2).$$

Notice that:

$$\|\boldsymbol{\Lambda}_k\|_2^2 \leqslant 2(\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2 + \|\boldsymbol{\lambda}\|_2^2).$$

Moreover by Hölder's inequality, Young's inequality, and the fact that $\mathbf{g}$ is continuous and $\mathcal{X}$ is compact, it holds:

$$\langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{g}(\boldsymbol{X}_k)\rangle \leqslant \|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2 \|\mathbf{g}(\boldsymbol{X}_k)\|_2 \leqslant \frac{\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2}{2} + \frac{\|\mathbf{g}(\boldsymbol{X}_k)\|_2^2}{2} \leqslant \frac{\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2}{2} + \frac{C}{2},$$

for a constant $C > 0$ independent of k. Combining all the results, it follows that:

$$\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2 \leqslant \|\boldsymbol{\Lambda}_0 - \boldsymbol{\lambda}\|_2^2 + \sum_{k=0}^{n-1} (\gamma_k + 8\gamma_k^2\alpha_k^2)(\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2 + C) + \sum_{k=0}^{n-1} (\gamma_k\alpha_k + 8\gamma_k^2\alpha_k^2)\|\boldsymbol{\lambda}\|_2^2 \\ + 4\sum_{k=0}^{n-1} \gamma_k^2 C_3^2,$$

and thus:

$$\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2 \leqslant \sum_{k=0}^{n-1} (\gamma_k\alpha_k + 8\gamma_k^2\alpha_k^2)\tilde{C} + \sum_{k=0}^{n-1} (\gamma_k + 8\gamma_k^2\alpha_k^2)\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2,$$

for a constant $\tilde{C} > 0$ independent of k. By setting:

$$y_n = \|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2, \quad f_n = \sum_{k=0}^{n-1} (\gamma_k\alpha_k + 8\gamma_k^2\alpha_k^2)\tilde{C}, \quad \text{and} \quad g_n = \gamma_n + 8\gamma_n^2\alpha_n^2,$$

we have by Gronwall's inequality as desired:

$$\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2 < \infty.$$

4.9.4. Proof of Theorem 4.14

In order to show Proposition 4.21, we need to give a high probability bound for both sums S_n and R_n under the light tail assumption of the summand (see Assumption 4.5).

In order to give a high probability bound for S_n, we introduce the following notion:

Definition 4.2 (Subgaussian Martingale): *Let (M_n) be a $\mathbb{F}$-martingale. (M_n) is called subgaussian if the increments $\Delta M_n := M_n - M_{n-1}$ is subgaussian given $\mathcal{F}_{t-1}$, i.e.*

for all n, there exists a constant $C > 0$. s.t.:

$$\mathbb{E}_{n-1}[\exp(\varDelta M_n^2/C^2)] \leqslant 2.$$

The infimum of all $C > 0$ s.t. above inequality holds is denoted by $\|\varDelta M_n\|_{\psi_2,n-1}$. We define the process:

$$\langle M\rangle_n^{\psi_2} := \sum_{k=1}^{n} \|\varDelta M_k\|_{\psi_2,k-1}^2, \quad n \in \mathbb{N}$$

It is immediate to see that (S_n) with $S_0 = 0$ is a subgaussian martingale. For a subgaussian martingale one can give a subgaussian concentration inequality:

Proposition 4.20 (Stochastic Exponential of Subgaussian Martingale): *Let (M_n) be a subgaussian martingale. Then:*

$$\mathbb{P}(M_n - M_0 \geqslant \epsilon,\ \langle M\rangle_n^{\psi_2} \leqslant r) \leqslant \exp(-\frac{\epsilon^2}{2C_{\psi_2} r}), \tag{4.62}$$

where $C_{\psi_2} = 2/e$ is an absolute constant.

Proof: Define the process:

$$\mathcal{E}(M,\lambda)_n^{\psi_2} := \exp(\lambda(M_n - M_0) - \lambda^2 C_{\psi_2} \langle M\rangle_n^{\psi_2}), \quad n \in \mathbb{N}_0$$

It holds:

$$\mathbb{E}_{n-1}[\mathcal{E}(M,\lambda)_n^{\psi_2}] = \mathcal{E}(M,\lambda)_{n-1}^{\psi_2} \cdot \mathbb{E}_{n-1}[\exp(\lambda(M_n - M_{n-1}) - \lambda^2 C_{\psi_2} \|M_n - M_{n-1}\|_{\psi_2,n}^2)].$$

Since $\mathbb{E}_{n-1}[M_n - M_{n-1}] = 0$, it follows from the subgaussian property see e.g. [117]:

$$\mathbb{E}_{t-1}[\exp(M_n - M_{n-1})] \leqslant \exp(C_{\psi_2} \|M_n - M_{n-1}\|_{\psi_2,n-1}^2),$$

Thus we obtain that $\mathcal{E}(M,\lambda)_n^{\psi_2}$, $n \in \mathbb{N}$, is a supermartingale

Now consider the event:

$$A_n := \left\{ M_n - M_0 \geqslant \epsilon,\ \langle M\rangle_n^{\psi_2} \leqslant r \right\}.$$

Markov inequality asserts that for $\lambda \geqslant 0$:

$$\begin{aligned}\mathbb{P}(A_n) &\leqslant \mathbb{E}\left[\exp\left(\frac{\lambda}{2}(M_n - M_0) - \frac{\lambda}{2}\epsilon\right)\mathbb{1}_{A_n}\right] \\ &\leqslant \mathbb{E}\left[\sqrt{\mathcal{E}(M,\lambda)_n^{\psi_2}}\exp\left(\frac{\lambda^2}{2}C_{\psi_2}\langle M\rangle_n^{\psi_2} - \frac{\lambda}{2}\epsilon\right)\mathbb{1}_{A_n}\right] \\ &\leqslant \mathbb{E}\left[\sqrt{\mathcal{E}(M,\lambda)_n^{\psi_2}}\exp\left(\frac{\lambda^2}{2}C_{\psi_2}r - \frac{\lambda}{2}\epsilon\right)\mathbb{1}_{A_n}\right],\end{aligned}$$

where the last inequality follows since on A_n, $\langle M\rangle_n^{\psi_2}$ is upper bounded by r. Cauchy-Schwarz inequality and the fact that $(\mathcal{E}(M,\lambda)_n^{\psi_2})_n$ is a supermartingale, give:

$$\mathbb{P}(A_n) \leqslant \mathbb{E}\left[\mathcal{E}(M,\lambda)_n^{\psi_2}\right]^{\frac{1}{2}}\mathbb{E}\left[\exp\left(\lambda^2 C_{\psi_2}r - \lambda\epsilon\right)\mathbb{1}_{A_n}\right]^{\frac{1}{2}} \leqslant \mathbb{E}\left[\exp\left(\lambda^2 C_{\psi_2}r - \lambda\epsilon\right)\mathbb{1}_{A_n}\right]^{\frac{1}{2}}.$$

Setting the optimal choice :

$$\lambda = \frac{\epsilon}{2C_{\psi_2}}$$

above, we have:

$$\mathbb{P}(A_n) \leqslant \exp\left(-\frac{\epsilon^2}{4C_{\psi_2}r}\right)\sqrt{\mathbb{P}(A_n)}.$$

By dividing both sides with $\sqrt{\mathbb{P}(A_n)}$, we obtain the desired statement ■

Definition 4.3 (Subexponential Submartingale): *Let (M_n) be a $\mathbb{F}$-submartingale. (M_n) is called subexponential if the increments $\Delta M_n := M_n - M_{n-1}$ is subexponential given $\mathcal{F}_{t-1}$, i.e. for all n, there exists a constant $C > 0$. s.t.:*

$$\mathbb{E}_{n-1}[\exp(|\Delta M_n|/C)] \leqslant 2.$$

The infimum of all $C > 0$ s.t. above inequality holds is denoted by $\|\Delta M_n\|_{\psi_1,n-1}$. We define the process:

$$\langle M\rangle_n^{\psi_1} := \sum_{k=1}^{n}\|\Delta M_k\|_{\psi_1,k-1}^2,$$

It is immediate to see that (R_n) with $R_0 = 0$ is a subexponential submartingale. For a subexponential submartingale one can give the following concentration inequality:

Proposition 4.21: *Let $(M_n)_n$ be a subexponential submartingale. If a.s. for all $n \geqslant 0$ there exists $c_n > 0$ s.t.:*

$$C_{\psi_1}\|M_n - M_{n-1}\|_{\psi_1,n-1} \leqslant c_n, \tag{4.63}$$

It holds:

$$\mathbb{P}\left(\sum_{k=1}^{n}|M_k - M_{k-1}| \geqslant C_{\psi_1}\langle M\rangle_n^{\psi_1} + \epsilon\right) \leqslant \exp\left(-\frac{\epsilon}{C_{\psi_1}(\sup_{k\in[n]_0} c_k)}\right), \tag{4.64}$$

where $C_{\psi_1} := 4e^{1+\frac{1}{e}}$ is an absolute constant.

Proof: Define the process:

$$\mathcal{E}(\lambda, M)_n^{\psi_1} := \exp(\sum_{k=1}^{n} \lambda \left|M_k - M_{k-1}\right| - \lambda C_{\psi_1} \left\langle M \right\rangle_n^{\psi_1}), \quad n \geqslant 0,$$

and let be:

$$0 \leqslant \lambda \leqslant \frac{1}{C_{\psi_1} \sup_n c_n} \tag{4.65}$$

It holds:

$$\mathbb{E}_{n-1}[\mathcal{E}(M,\lambda)_n^{\psi_1}] = \mathcal{E}(M,\lambda)_{n-1}^{\psi_1} \cdot \mathbb{E}_{n-1}[\exp(\lambda \left|M_n - M_{n-1}\right| - C_{\psi_1} \lambda \|M_n - M_{n-1}\|_{\psi_1,n})].$$

By subexponential property [117], we have for λ satisfying (4.65):

$$\mathbb{E}_{t-1}[\exp(\lambda \left|M_t - M_{t-1}\right| - \lambda C_{\psi_1} \|M_t - M_{t-1}\|_{\psi_1,t})] \leqslant 1.$$

This shows that $\mathcal{E}(M,\lambda)_n^{\psi_1}$ is a supermartingale. Now by Markov's inequality and martingale property of $\mathcal{E}(M,\lambda)_n^{\psi_1}$, it holds:

$$\mathbb{P}(\sum_{k=1}^{n} \left|M_k - M_{k-1}\right| \geqslant C_{\psi_1} \left\langle M \right\rangle_n^{\psi_1} + \epsilon) \leqslant \inf_{\boldsymbol{\lambda} \in [0, 1/C_{\psi_1} \sup_n c_n)} \exp(-\lambda\epsilon).$$

Setting the optimal λ we obtain (4.64). ∎

Combaining both proposition, we obtain Theorem 4.21:

Proof (Proof of Theorem 4.14): We set $S_0 = 0$.

$$\|\Delta S_n\|_{\psi_2,n-1} \leqslant C_{\mathcal{X}} \gamma_{n-1} \|\|\xi_n\|_*\|_{\psi_2,n-1} \leqslant C_{\mathcal{X}} \gamma_{n-1} \sigma_{n-1}$$

Thus:

$$\left\langle S \right\rangle_n^{\psi_2} = \sum_{k=1}^{n} \|\Delta S_n\|_{\psi_2,n-1}^2 \leqslant C_{\mathcal{X}}^2 \sum_{k=0}^{n-1} \gamma_k^2 \sigma_k^2 \quad \text{a.s.}$$

By Proposition 4.20 it holds with probability at least $1-\delta$:

$$S_n \leqslant \sqrt{2\ln(1/\delta) C_{\psi_2} C_{\mathcal{X}}^2 \sum_{k=0}^{n-1} \gamma_k^2 \sigma_k^2}$$

It holds:

$$\|\Delta R_n\|_{\psi_1,n-1} = \gamma_{n-1}^2 \|\|\xi_n\|_*\|_{\psi_1,n-1} \leqslant \gamma_{n-1}^2 \sigma_{n-1}^2$$

Proposition 4.21 asserts that with probability at least $1-\delta$, we have:

$$R_n \leqslant C_{\psi_1} \sum_{k=0}^{n-1} \gamma_k^2 \sigma_k^2 + \ln(1/\delta) C_{\psi_1} \sup_{k \in [n]} \sigma_{k-1}^2. \qquad \blacksquare$$

4.10. Choices of Step size and Augmentation Sequence

4.10.1. Detailed explanation of Remark 15

Set $\boldsymbol{Y}_0 = 0$ and $\boldsymbol{\Lambda}_0 = 0$, it follows that:

$$\tilde{F}(\boldsymbol{z}_*, \boldsymbol{Z}_0) \leqslant \frac{\sum_{i=1}^N \Delta\psi_i(\mathcal{X}_i)}{N} + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2}, \quad \text{where } \Delta\psi_i(\mathcal{X}_i) := \max_{\mathcal{X}_i} \psi_i - \min_{\mathcal{X}_i} \psi_i.$$

Let us denote:

$$\overline{\Delta\psi(\mathcal{X})} := \frac{\sum_{i=1}^N \Delta\psi_i(\mathcal{X}_i)}{N}$$

In case that we have constant step size $\gamma_n = \gamma$ Theorem 4.14 asserts:

$$\mathbb{E}[\overline{\Xi}_n^\gamma] \leqslant \left[\frac{\overline{\Delta\psi(\mathcal{X})}}{\gamma n} + (\tilde{C}_1 + \frac{2G^2}{K})\gamma \right] + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} \left[\frac{1}{\gamma n} + \alpha\gamma \right].$$

If we choose:

$$\gamma = \sqrt{\frac{\overline{\Delta\psi(\mathcal{X})}}{(\tilde{C}_1 + \frac{2G^2}{K})n}},$$

for large enough n s.t. (4.40) holds, then:

$$\mathbb{E}[\overline{\Xi}_n^\gamma] \leqslant \sqrt{\overline{\Delta\psi(\mathcal{X})}(\tilde{C}_1 + \frac{2G^2}{K})} \frac{2}{\sqrt{n}} + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2\sqrt{n}} \left[\sqrt{\frac{(\tilde{C}_1 + \frac{2G^2}{K})}{\overline{\Delta\psi(\mathcal{X})}}} + \alpha \sqrt{\frac{\overline{\Delta\psi(\mathcal{X})}}{(\tilde{C}_1 + \frac{2G^2}{K})}} \right].$$

This shows that the complexity bound of order $\mathcal{O}(Gn^{-1/2})$ is achievable for fixed stepsize and fixed time horizon.

Now suppose that $\gamma_n = \gamma/(n+1)$, where $\gamma > 0$ fulfills the inequality (4.40). It follows from Theorem 4.14:

$$\mathbb{E}[\overline{\Xi}_n^\gamma] \leqslant \left[\frac{\overline{\Delta\psi(\mathcal{X})}}{\gamma} + (\tilde{C}_1 + \frac{2G^2}{K})\gamma + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} \left[\frac{1}{\gamma} + \alpha\gamma \right] \right] \mathcal{O}(n^{-1/2}\ln(n)).$$

This shows that the complexity bound of order $\mathcal{O}(G^2 \ln(n) n^{-1/2})$ is achievable for variable step size and augmentation sequence.

4.10.2. Detailed explanation of Remark 16

If we choose $\boldsymbol{Y}_0 = 0$ and $\boldsymbol{\Lambda}_0 = 0$, it follows that:

$$\tilde{F}(\boldsymbol{z}_*, \boldsymbol{Z}_0) \leqslant \frac{\sum_{i=1}^N \Delta\psi_i(\mathcal{X}_i)}{N} + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2}, \quad \text{where } \Delta\psi_i(\mathcal{X}_i) := \max_{\mathcal{X}_i} \psi_i - \min_{\mathcal{X}_i} \psi_i.$$

Let us denote:

$$\overline{\Delta\psi(\mathcal{X})} := \frac{\sum_{i=1}^N \Delta\psi_i(\mathcal{X}_i)}{N}$$

It holds:

$$\Xi_k^\gamma \leqslant \left[\frac{\overline{\Delta\psi(\mathcal{X})} + \ln(2/\delta)C_{\psi_1}\sigma^2}{\gamma n} + (\tilde{C}_1 + \sigma^2\frac{2C_{\psi_1}}{K})\gamma\right] + \frac{\sigma C_{\mathcal{X}}\sqrt{2\ln(2/\delta)C_{\psi_2}}}{\sqrt{n}} + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2}\left[\frac{1}{\gamma n} + \alpha\gamma\right].$$

So for a fixed $n \geqslant 0$, if we choose:

$$\gamma = \sqrt{\frac{\overline{\Delta\psi(\mathcal{X})} + \ln(2/\delta)C_{\psi_1}\sigma^2}{(\tilde{C}_1 + \sigma^2\frac{2C_{\psi_1}}{K})n}},$$

then with probability at least $1-\delta$:

$$\Xi_k^\gamma \leqslant 2\sqrt{\frac{(\overline{\Delta\psi(\mathcal{X})} + \ln(2/\delta)C_{\psi_1}\sigma^2)(\tilde{C}_1 + \sigma^2\frac{2C_{\psi_1}}{K})}{n}} + \frac{\sigma C_{\mathcal{X}}\sqrt{2\ln(2/\delta)C_{\psi_2}}}{\sqrt{n}}$$
$$+ \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2\sqrt{n}}\left[\sqrt{\frac{(\tilde{C}_1 + \sigma^2\frac{2C_{\psi_1}}{K})}{\overline{\Delta\psi(\mathcal{X})} + \ln(2/\delta)C_{\psi_1}\sigma^2}} + \alpha\sqrt{\frac{\overline{\Delta\psi(\mathcal{X})} + \ln(2/\delta)C_{\psi_1}\sigma^2}{(\tilde{C}_1 + \sigma^2\frac{2C_{\psi_1}}{K})}}\right].$$

Thus with probability at least $1-\delta$, we have the complexity bound:

$$\Xi_k^\gamma \leqslant \mathcal{O}\left(\sqrt{\frac{\ln(2/\delta)\sigma^2}{n}}\right).$$

Equivalently, to get the bound $\overline{\Xi}_n^\gamma \leqslant \epsilon$ with probability at least $1-\delta$, we need:

$$n \geqslant \mathcal{O}\left(\frac{\ln(2/\delta)\sigma^2}{\epsilon^2}\right)$$

Now suppose that $\gamma_n = \gamma/(n+1)$, where $\gamma > 0$ fulfills the inequality (4.40). Since in this case $\sum_{k=0}^{n-1}\gamma_k \lesssim n^{-1/2}$ and $\sum_{k=0}^{n-1}\gamma_k \lesssim ln(n)$, ot holds:

$$\mathbb{E}[\Xi_n^\gamma] \leqslant \left[\frac{\overline{\Delta\psi(\mathcal{X})} + \ln(2/\delta)C_{\psi_1}\sigma^2}{\gamma} + (\tilde{C}_1 + \frac{2\sigma^2C_{\psi_1}}{K})\gamma + \frac{\sigma C_{\mathcal{X}}\sqrt{2\ln(2/\delta)C_{\psi_2}}}{\sqrt{n}} + \frac{\|\boldsymbol{\lambda}_*\|_2^2}{2}\left[\frac{1}{\gamma} + \alpha\gamma\right]\right]$$
$$\cdot\,\mathcal{O}\left(\frac{\ln(n)}{\sqrt{n}}\right).$$

Thus with probability at least $1 - \delta$, we have the complexity bound:

$$\Xi_n^{\gamma} \leqslant \mathcal{O}\left(\frac{\ln(2/\delta)\sigma^2 \ln(n)}{\sqrt{n}}\right).$$

5. Impact of Agents' Price Sensitivity on the Resource Sustainable Pricing

5.1. Introduction

In this chapter, we continue the discussion made in the last chapter, where we have designed an incentive-based control mechanism based on resource pricing aiming to foster resource-sustainable behavior in a system of competitive online learning agents. There, we have seen, by charging each agents additional cost for the amount of utilization of resources with prices proportional to the capacity violation of the resources, can lead to the decay of the average of the resources' capacity violations.

This observation made in the last chapter enlightens the following role of the price: It provides information about the degree of the scarcity of a resource, measured by the amount of the violation of their capacity constraints. This role of the price is also known in economic theory. One fundamental principle in this direction is the so-called scarcity principle (see e.g. [118]), which explains the dynamic of the price of an economic good as the result of a mismatch between supply and demand. This principle asserts in particular that the price of a scarce good increases until the supply and demand reach an equilibrium.

So naively, one may think, that since the prices are linked to the actual congestion state and are the only provider of amount of resource constraint violations, the agents should give up their personal preferences and choose their decision according to the actual price of the resources. Accordingly, the question which we ask in this chapter is the following:

If the learning agents would be more sensitive to the prices as to their own preferences, does the resource-centric pricing lead to a more resource sustainable behaviour in a competitive system?

In this chapter, we show that, to a certain degree, price sensitiveness can foster sustainable behaviour. However, a high price sensitiveness can be in contrary to the goal of establishing that behaviour. Intuitive reasoning of the latter claim is that high price sensitiveness may cause herding behaviour in the utilization of the resources: the agents tend to use resources which has the lowest price, equally given for all agents, and not the resources which provide them the highest benefit.

This chapter is organized as follows:

- In Section 5.2 we extend Algorithm 4 by introducing an additional parameter called sensitivity parameter. With the sensitivity parameter, one can increases the importance of the price in the agents' decision making. There, we also provide an informal discussion about how this parameter influences the populations' behaviour.

- Subsequently in Section 5.3, we provide for a certain sensitivity parameter choice a guarantee for resource sustainable behaviour. Although the guarantee is comparable to that for Algorithm 3, it is more convenient.

- Finally in Section 5.4 we provide some numerical simulations in order to support both the informal discussions given in Section 5.2 and the theoretical guarantee given in the Section 5.3.

5.2. Basic Setting and Price Mechanism

Non-Cooperative Game with Coupled Resource Constraints Here, we consider the similar basic setting as in the previous chapter, i.e., the setting of non cooperative game (See Section 2.2) with coupled constraints (See Subsection 4.2.1), with the difference that we assume that the underlying coupled constraints are affine, i.e., we consider the function $\mathbf{g}$ specified in Assumption 4.1 takes the form

$$\mathbf{g}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x} - b,$$

where:

$$\mathbb{R}^{R\times\sum_{i=1}^{N} D_i} \ni \mathbf{A} := [\mathbf{A}_1, \ldots, \mathbf{A}_N] \quad \text{and} \quad b \in \mathbb{R}^R,$$

with for any $i \in [N]$:

$$\mathbf{A}_i \in \mathbb{R}^{R\times D_i}$$

is the matrix which is available only to the agent i. This specification of the resource constraints is for sake of simplicity of the arguments. However for the general setting, one can straightforwardly extend the given arguments. Furthermore, we choose the affine constraints to exploit the decentralization aspect discussed briefly in Remark 5.

Agents' Model and Pricing Mechanism Same as in the previous chapter, we assume that the agents are online learner applying the online mirror descent algorithm. In order to give the agents incentives for sustainable use of resources, our advice is to charge each agent additional cost for the amount of utilization of resources related to her action. Specifically, consider a time slot t, agent i is obligate to pay $\tilde{\boldsymbol{\Lambda}}_t^{\mathrm{T}}\mathbf{A}_i\boldsymbol{X}_{t+1}^{(i)}$ for a possible future action $\boldsymbol{X}_{t+1}^{(i)} \in \mathcal{X}_i$, where $\tilde{\boldsymbol{\Lambda}}_t$ is a vector specifying the price of each resource at time t. So at time t, the utility function of agent i becomes $\mathrm{u}_t^{(i)}(\cdot) + \tilde{\boldsymbol{\Lambda}}_t^T\mathbf{A}_i(\cdot)$, and correspondingly assuming that the price information is not noisy, the gradient update (4.3) turns to:

$$\boldsymbol{Y}_{t+1}^{(i)} = \boldsymbol{Y}_t^{(i)} + \gamma_t(\hat{\mathbf{v}}_t^{(i)} + \mathbf{A}_i^{\mathrm{T}} \tilde{\boldsymbol{\Lambda}}_t). \tag{5.1}$$

The update of each entry of the price vector $\tilde{\boldsymbol{\Lambda}}_{t+1}$ is done by each of the resources separately proportional to their own congestion state. The specific mechanism is provided in Algorithm 4.

Algorithm 4

Require: Horizon length T,

- For each $t \in [T-1]_0$: agents' learning rate $\gamma_\tau > 0$, resources' learning rate $\zeta_k \in (0,1)$, price progressivity $\eta_k \geqslant 0$, price sensitivity $\beta_k \geqslant 0$,
- Initialization: Score vectors $\boldsymbol{Y}_0^{(i)} \in \mathbb{R}^{D_i}$, $i \in [N]$, prices $\boldsymbol{\Lambda}_0^r \in \mathbb{R}_{\geqslant 0}$, $r \in [R]$.

//Mechanism

for $t = 1, 2, \ldots, T$ **do**
  Every agent $i \in [N]$ mutually play $\boldsymbol{X}_t^{(i)} \leftarrow \boldsymbol{\Phi}_i(\boldsymbol{Y}_t^{(i)})$
  //Decision making via Online Learning
  for every player $i \in [N]$ **do**
    Observe noisy gradient utility feedback (7.3) and update the score vector $\boldsymbol{Y}_{t+1}^{(i)}$ via (5.1)
    Query the prices $\tilde{\boldsymbol{\Lambda}}_t^r$ from the resources $r \in [R]$
  end for
  //Pricing
  for every resource $r \in [R]$ **do**
    Check the actual own congestion state: $\boldsymbol{\phi}_t^r = \boldsymbol{\phi}^r(\boldsymbol{X}_t) = [\mathbf{A}\boldsymbol{X}_t - \boldsymbol{b}]_r$
    Update the price:

$$\boldsymbol{\Lambda}_{t+1}^r \leftarrow [(1-\eta_t)\boldsymbol{\Lambda}_t^r + \zeta_t \boldsymbol{\phi}_t^r]_+ \quad \text{and} \quad \tilde{\boldsymbol{\Lambda}}_{t+1}^r \leftarrow \beta_t \cdot \boldsymbol{\Lambda}_{t+1}^r$$

  end for
end for

Pricing Parameters Choice – Gedankenexperiment The parameter β_k specifies to what extent the price of a resource should be considered in the decision-making process of the agents. In order to understand the effect of this parameter to the population dynamic, let us consider the extreme cases $\beta_k = 0$ and high $\beta_k > 0$. With $\beta_k = 0$, the population dynamic described in Algorithm 4 turns to (4.3). As already discussed in the previous chapter extensively, we cannot expect decaying resource congestion. An intuitive reasoning for this occurance is that if $\beta_k = 0$, the corresponding population dynamic potentially converges to the stable set of variational Nash Equilibrium of the non-cooperative game, which generally does not satisfy the coupled constraints. Now, if $\beta_k > 0$ is high, the agents tend to take the action with cheapest cost. Since the price of a resource is proportional to its congestion state, all agents might at worst (e.g. in the case

$N = D$ and $\mathbf{A}_i = \mathbf{I}_N$, where an action corresponds directly to resource utilization choice) fully consume a single resource with the lowest congestion and cause therefore the latter's price and load to rise dramatically. Subsequently in the next time slot, they will all mutually fully utilized another less congested and cheaper resource causing its price and congestion to rise dramatically. This procedure will repeat, cause agents' consumption choice bounces at worst from a single resource to another one, and meanwhile violation of resource capacity constraints. This gedankenexperiment asserts in particular that high prices and thus high degree of control, in contrary to the intuition, does in general not support sustainable behaviour. Rather, one should allow for the latter's sake to a certain degree egoistic behaviour of the agents.

The parameter η_k specifies the strength of the dependency of the price update on the previous price. $\eta_t = 1$ corresponds to the extreme case where the price update only based on the actual congestion state $\boldsymbol{\phi}_t^r$. We expect $\eta_k = 1$ is not a good choice since it ignores the price dynamic and correspondingly the agents' consumption behaviour implicitly described therein. A problem which might occur with the extreme case $\eta_k = 0$ is the rapid increase of the prices causing the price update insensitive against changes in the congestion state of the resources.

5.3. Non-asymptotic Guarantee of the Price Mechanism

In this section we provide a theoretical analysis of the price mechanism provided in Algorithm 4. Our emphasize is on the degree of its contribution to the resource-aware consumption behaviour of the agents, which we measured by the (time) *average of the norm of the clipped cumulative violation of constraints (ANCCVC)*:

$$\mathrm{ANCCVC}_t := \frac{\mathbb{E}\left[\| \left[\sum_{t=0}^{t-1}(\mathbf{A}\boldsymbol{X}_t - \boldsymbol{b})\right]_+ \|_2\right]}{t}, \quad t \in \mathbb{N}.$$

ANCCVC gives in particular an estimate for the time average congestion state of the resource since:

$$\mathrm{ANCCVC}_t \geqslant \sum_{k=0}^{t-1} \frac{\boldsymbol{\phi}_r(\boldsymbol{X}_k)}{t}$$

for all $r \in [R]$.

Throughout, C_1, C_2, C_3 denote non-negative constants fulfilling for all $x \in \mathcal{X}$ and $\boldsymbol{\lambda} \in \mathbb{R}^M_{\geqslant 0}$:

$$\|\mathbf{A}^T\boldsymbol{\lambda}\|_* \leqslant C_1\|\boldsymbol{\lambda}\|_2, \ \|\mathbf{v}(\boldsymbol{x})\|_* \leqslant C_2, \|g(\boldsymbol{x})\|_2 \leqslant C_3, \tag{5.2}$$

which clearly exists by our assumptions on u and $\mathcal{X}$. Our main result is the following:

Theorem 5.1: *Given a horizon length $n \in \mathbb{N}$ and learning rate $\gamma_k > 0$, $k \in [n-1]_0$. Set* *the extrinsic price sensitivity of the agents and the resources'learning rate as:*

$$\beta_k = 2 \quad \textit{and} \quad \zeta_k = \gamma_k, \quad \forall k \in [t-1]_0,$$

and suppose that for all $k \in [t-1]_0$ the agents' learning rate and the price progressivity fulfills:

$$\eta_k^2 - \frac{\eta_k}{2} + \frac{2\gamma_k^2 C_1^2}{K} \leqslant 0, \quad \forall k \in [n-1]_0 \qquad \textit{(Trackability Condition (TC))}, \tag{5.3}$$

Then it holds for $\boldsymbol{\Lambda}_0 = 0$ and $\boldsymbol{Y}_0 = 0$:

$$\mathbb{E}\left[\|[\sum_{k=0}^{n-1} \gamma_k (\mathbf{A}\boldsymbol{X}_k - \boldsymbol{b})]_+\|_2^2 \right] \leqslant 2\overline{\eta}_n \left(\Delta\psi + \tilde{C}_1 \sum_{k=0}^{n-1} \gamma_k^2 \right) + \overline{\eta}_n^2 \left(\|\boldsymbol{\lambda}_*\|_2^2 + \frac{4}{K} \sum_{k=1}^{n} \gamma_{k-1}^2 \mathbb{E}[\|\boldsymbol{M}_k\|_*^2] \right), \tag{5.4}$$

where:

$$\overline{\eta}_t := \sum_{k=0}^{t-1} \eta_k + 1, \; \tilde{C}_1 := 2\left(\frac{C_2^2}{K} + C_3^2 \right), \; \Delta\psi = \sum_{i=1}^{N} \left(\max_{\mathcal{X}_i} \psi_i - \min_{\mathcal{X}_i} \psi_i \right), \; K := \min_i K_i,$$

and where $\boldsymbol{\lambda}^* \in \mathbb{R}^R_{\geqslant 0}$ fulfills:

$$(\boldsymbol{x}_*, \boldsymbol{\lambda}) \in \mathrm{SOL}(\mathcal{X} \times \mathbb{R}^R_{\geqslant 0}, \tilde{\mathbf{v}}), \quad \textit{for an } \boldsymbol{x}_* \in \mathcal{X}.$$

Remark 17: In contrast to the approach made in Chapter 4, the proof of this theorem is based on the analysis of the dynamic of the energy function:

$$\mathcal{E}_t((\boldsymbol{x}, \boldsymbol{\lambda}), \tilde{\boldsymbol{\lambda}}) := \mathcal{E}_t^{(1)}((\boldsymbol{x}, \boldsymbol{\lambda})) + \mathcal{E}_t^{(2)}(\tilde{\boldsymbol{\lambda}}),$$

where:

$$\mathcal{E}_k^{(1)}((\boldsymbol{x}, \boldsymbol{\lambda})) := \mathrm{F}(\boldsymbol{x}, \boldsymbol{Y}_k) + \frac{\|\boldsymbol{\Lambda}_k - \boldsymbol{\lambda}\|_2^2}{2}, \quad \mathcal{E}_k^{(2)}(\tilde{\boldsymbol{\lambda}}) := \frac{\|\boldsymbol{\Lambda}_k - \tilde{\boldsymbol{\lambda}}\|_2^2}{2}, \quad \boldsymbol{x} \in \mathcal{X}, \quad \text{and} \quad \boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}} \in \mathbb{R}^R.$$

Moreover, to generate this result we do not analyze the dynamic of the price update.

Now, we are ready to give the proof of Theorem

Proof (Proof of Theorem 5.1): For all $k \in \mathbb{N}$ and $\boldsymbol{x} \in \mathcal{X}$, it holds for $\mathcal{V}_n^{(1)}(\boldsymbol{x}) := \mathrm{F}(\boldsymbol{x}, \boldsymbol{Y}_t) - \mathrm{F}(\boldsymbol{x}, Y_0)$:

$$\begin{aligned} \mathcal{V}_n^{(1)}(\boldsymbol{x}) \leqslant & \sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{X}_k - \boldsymbol{x}, \mathbf{v}(\boldsymbol{X}_k) \rangle - \sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{X}_k - \boldsymbol{x}, \beta_k \mathbf{A}^{\mathrm{T}} \boldsymbol{\Lambda}_k \rangle \\ & + \sum_{k=0}^{n-1} \frac{\beta_k^2 \gamma_k^2 C_1^2}{K} \|\boldsymbol{\Lambda}_k\|_2^2 + S_n(\boldsymbol{x}) + \frac{2}{K} R_n + \frac{2C_2^2 \sum_{k=0}^{n-1} \gamma_k^2}{K}, \end{aligned}$$

where C_1 and C_2 are given in (5.2), and where:

$$S_n(\boldsymbol{x}) := \sum_{k=0}^{n-1} \gamma_k \langle \boldsymbol{X}_k - \boldsymbol{x}, \boldsymbol{M}_{k+1} \rangle, \;\; R_n := \sum_{k=0}^{n-1} \gamma_k^2 \|\boldsymbol{M}_{k+1}\|_*^2.$$

Now, we analyze the dynamic of the energy function $\|\boldsymbol{\Lambda}_n - \boldsymbol{\lambda}\|_2^2/2$. it holds for $\mathcal{V}_t^{(2)}(\boldsymbol{\lambda}) := (\|\boldsymbol{\Lambda}_t - \boldsymbol{\lambda}\|_2^2 - \|\boldsymbol{\Lambda}_0 - \boldsymbol{\lambda}\|_2^2)/2$:

$$\mathcal{V}_t^{(2)}(\boldsymbol{\lambda}) \leqslant \sum_{k=0}^{n-1} \zeta_k \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle + \sum_{k=1}^{n-1} \frac{(2\eta_k^2 - \eta_k)}{2} \|\boldsymbol{\Lambda}_k\|_2^2 + \|\boldsymbol{\lambda}\|_2^2 \sum_{k=0}^{n-1} \frac{\eta_k}{2} + C_3^2 \sum_{k=0}^{n-1} \zeta_k^2. \tag{5.5}$$

Combining the previous bounds for $\mathcal{V}_t^{(1)}(\boldsymbol{x})$ and $\mathcal{V}_t^{(2)}(\boldsymbol{\lambda})$, it yields for $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{\lambda}) \in \mathcal{X} \times \mathbb{R}_{\geqslant 0}^R$:

$$\begin{aligned}
\mathcal{V}_t(\boldsymbol{z}) \leqslant & - \sum_{k=0}^{t-1} \gamma_k \Xi_k(\boldsymbol{z}) + \sum_{k=0}^{t-1} \gamma_k (1-\beta_k) \langle \boldsymbol{X}_\tau - \boldsymbol{x}, \mathbf{A}^{\mathrm{T}} \boldsymbol{\Lambda}_k \rangle + \frac{2C_2^2}{K} \sum_{k=1}^{t-1} \gamma_k^2 + C_3^2 \sum_{k=1}^{t-1} \zeta_k^2 \\
& + \|\boldsymbol{\lambda}\|_2^2 \sum_{k=0}^{t-1} \frac{\eta_k}{2} + \sum_{k=0}^{t-1} \left(\eta_k^2 - \frac{\eta_k}{2} + \frac{\beta_k^2 \gamma_k^2 C_1^2}{K} \right) \|\boldsymbol{\Lambda}_k\|_2^2 + S_t(\boldsymbol{x}) + \frac{2}{K} R_t \\
& + \sum_{k=0}^{t-1} (\gamma_k - \zeta_k) \langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle,
\end{aligned}$$

where $\mathcal{V}_t(\boldsymbol{z}) = \mathcal{V}_t^{(1)}(\boldsymbol{x}) + \mathcal{V}_t^{(2)}(\boldsymbol{\lambda})$, and where:

$$\Xi_t(\boldsymbol{z}) := \langle \boldsymbol{z} - \boldsymbol{Z}_t, \tilde{\mathbf{v}}(Z_t) \rangle.$$

By straightforward computation one can show that $\mathbf{v}$ monotone asserts $\tilde{\mathbf{v}}$ is monotone. Thus it holds:

$$\Xi_t(\boldsymbol{z}) \geqslant \langle \boldsymbol{z}_* - \boldsymbol{Z}_t, \tilde{\mathbf{v}}(\boldsymbol{z}_*) \rangle \geqslant 0, \quad \forall z_* \in \mathrm{SOL}(\mathcal{X} \times \mathbb{R}_{\geqslant 0}^R, \tilde{\mathbf{v}}).$$

This and the choice $\zeta_k = \gamma_k$ yields for $z_* = (\boldsymbol{x}_*, \boldsymbol{\lambda}_*) \in \mathrm{SOL}(\mathcal{X} \times \mathbb{R}_{\geqslant 0}^R, \tilde{\mathbf{v}})$:

$$\begin{aligned}
\mathcal{V}_t(\boldsymbol{z}_*) \leqslant & \sum_{k=0}^{t-1} \gamma_k (1-\beta_k) \langle \boldsymbol{X}_k - \boldsymbol{x}_*, \mathbf{A}^{\mathrm{T}} \boldsymbol{\Lambda}_k \rangle + \frac{2C_2^2}{K} \sum_{k=1}^{t-1} \gamma_k^2 + C_3^2 \sum_{k=1}^{t-1} \zeta_k^2 \\
& + \|\boldsymbol{\lambda}_*\|_2^2 \sum_{k=0}^{t-1} \frac{\eta_k}{2} + \sum_{k=0}^{t-1} \left(\eta_k^2 - \frac{\eta_k}{2} + \frac{\beta_k^2 \gamma_k^2 C_1^2}{K} \right) \|\boldsymbol{\Lambda}_k\|_2^2 + S_t(\boldsymbol{x}_*) + \frac{2}{K} R_t
\end{aligned} \tag{5.6}$$

To eliminate the first summand in above bound, we continue:

Lemma 5.2: *It holds for all $\boldsymbol{\lambda} \geqslant 0$:*

$$\langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle \leqslant \langle \boldsymbol{X}_k - \tilde{\boldsymbol{x}}, \mathbf{A}^{\mathrm{T}}\boldsymbol{\Lambda}_k \rangle - \langle \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle,$$

where $\tilde{\boldsymbol{x}} \in \mathcal{Q}$ arbitrary.

Proof (Proof of Lemma 5.2): we have for any $x \in \mathcal{X}$:

$$\begin{aligned}
\langle \boldsymbol{\Lambda}_k - \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle &= \langle \boldsymbol{\Lambda}_k, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle - \langle \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle \\
&= \langle \boldsymbol{\Lambda}_k, \mathbf{A}\boldsymbol{X}_k - \mathbf{A}x \rangle + \langle \boldsymbol{\Lambda}_k, \mathbf{A}x - \boldsymbol{b} \rangle - \langle \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle \\
&= \langle \boldsymbol{X}_k - \boldsymbol{x}, \mathbf{A}^{\mathrm{T}}\boldsymbol{\Lambda}_k \rangle + \langle \boldsymbol{\Lambda}_k, \mathbf{A}x - \boldsymbol{b} \rangle - \langle \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle. \qquad (5.7)
\end{aligned}$$

Now, since $\boldsymbol{\Lambda}_k \geqslant 0$, it follows that $\langle \boldsymbol{\Lambda}_k, \mathbf{A}\tilde{\boldsymbol{x}} - \boldsymbol{b} \rangle \leqslant 0$, for $\tilde{\boldsymbol{x}} \in \mathcal{Q}$. Therefore, if we set $x = \tilde{\boldsymbol{x}}$ with $\boldsymbol{x}_* \in \mathcal{Q}$ in (5.7), we have from previous observation the desired statement. ■

Setting this observation into (5.5) and setting the choice $\zeta_k = \gamma_k$, it yields for any $\tilde{\boldsymbol{x}} \in \mathcal{Q}$:

$$\begin{aligned}
\mathcal{V}_t^{(2)}(\boldsymbol{\lambda}) \leqslant& \sum_{k=0}^{t-1} \gamma_k \langle \boldsymbol{X}_k - \tilde{\boldsymbol{x}}, \mathbf{A}^{\mathrm{T}}\boldsymbol{\Lambda}_k \rangle - \sum_{k=0}^{t-1} \gamma_k \langle \boldsymbol{\lambda}, \mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \rangle \\
&+ \sum_{k=1}^{t-1} \frac{(2\eta_k^2 - \eta_k)}{2} \|\boldsymbol{\Lambda}_k\|_2^2 + \|\boldsymbol{\lambda}\|_2^2 \sum_{k=0}^{t-1} \frac{\eta_k}{2} + C_3^2 \sum_{k=0}^{t-1} \gamma_k^2
\end{aligned}$$

For $z_* = (\boldsymbol{x}_*, \boldsymbol{\lambda}_*)$ $\boldsymbol{\lambda}_* \geqslant 0$, and for $\tilde{\boldsymbol{\lambda}} \geqslant 0$, it holds by combining above inequality (with $\tilde{\boldsymbol{x}} = \boldsymbol{x}_*$) and (5.6):

$$\begin{aligned}
\mathcal{V}_t(\boldsymbol{z}_*) + \mathcal{V}_t^{(2)}(\tilde{\boldsymbol{\lambda}}) \leqslant& -\left(\langle \tilde{\boldsymbol{\lambda}}, \sum_{k=0}^{t-1} \gamma_k \left(\mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \right) \rangle - \frac{\|\tilde{\boldsymbol{\lambda}}\|_2^2}{2} \sum_{k=0}^{t-1} \eta_k \right) \\
&+ \sum_{k=0}^{t-1} \gamma_k (2 - \beta_k) \langle \boldsymbol{X}_k - \boldsymbol{x}, \mathbf{A}^{\mathrm{T}}\boldsymbol{\Lambda}_k \rangle + \left(2\eta_k^2 - \eta_k + \frac{\beta_k^2 \gamma_k^2 C_1^2}{K} \right) \sum_{k=0}^{t-1} \|\boldsymbol{\Lambda}_k\|_2^2 \\
&+ 2\left(\frac{C_2^2}{K} + C_3^2 \right) \sum_{k=0}^{t-1} \gamma_k^2 + \|\boldsymbol{\lambda}_*\|_2^2 \sum_{k=0}^{t-1} \frac{\eta_k}{2} + S_t(\boldsymbol{x}_*) + \frac{2}{K} R_t.
\end{aligned}$$

The second summand is eliminated by the choice $\beta_k = 2$ and the third summand by the trackability condition. Thus it follows:

$$\begin{aligned}
\mathcal{V}_t(\boldsymbol{z}_*) + \mathcal{V}_t^{(2)}(\tilde{\boldsymbol{\lambda}}) \leqslant& -\left(\langle \tilde{\boldsymbol{\lambda}}, \sum_{k=0}^{t-1} \gamma_k \left(\mathbf{A}\boldsymbol{X}_k - \boldsymbol{b} \right) \rangle - \frac{\|\tilde{\boldsymbol{\lambda}}\|_2^2}{2} \sum_{k=0}^{t-1} \eta_k \right) + 2\left(\frac{C_2^2}{K} + C_3^2 \right) \sum_{k=0}^{t-1} \gamma_k^2 \\
&+ \|\boldsymbol{\lambda}_*\|_2^2 \sum_{k=0}^{t-1} \frac{\eta_k}{2} + S_t(\boldsymbol{x}_*) + \frac{2}{K} R_t.
\end{aligned} \qquad (5.8)$$

Now, since $\boldsymbol{\Lambda}_0 = 0$, one sees that $\mathcal{V}_t^{(2)}(\tilde{\boldsymbol{\lambda}}) \geqslant -\|\tilde{\boldsymbol{\lambda}}\|_2^2/2$. Moreover since $\boldsymbol{\Lambda}_0 = 0$ and $Y_0 = 0$, we have $\mathcal{V}_t(\boldsymbol{z}_*) \geqslant -\Delta\psi(\mathcal{X}) - (\|\boldsymbol{\lambda}_*\|_2^2/2)$. Combining those observations with (5.8), we obtain:

$$\left[\langle\tilde{\boldsymbol{\lambda}}, \sum_{k=0}^{t-1}\gamma_k(\mathbf{A}\boldsymbol{X}_k - \boldsymbol{b})\rangle - \frac{\sum_{k=0}^{t-1}\eta_k + 1}{2}\|\tilde{\boldsymbol{\lambda}}\|_2^2\right] \leqslant \Delta\psi(\mathcal{X}) + 2\left(\frac{C_2^2}{K} + C_3^2\right)\sum_{k=1}^{t-1}\gamma_k^2 + \frac{(\sum_{k=0}^{t-1}\eta_k + 1)}{2}\|\boldsymbol{\lambda}_*\|_2^2 + S_t(\boldsymbol{x}_*) + \frac{2}{K}R_t. \tag{5.9}$$

Since:

$$\sup_{\tilde{\boldsymbol{\lambda}}\geqslant 0}\left(\langle\tilde{\boldsymbol{\lambda}}, \sum_{k=0}^{t-1}\gamma_k\left(\mathbf{A}\boldsymbol{X}_k - \boldsymbol{b}\right)\rangle - \frac{\sum_{k=0}^{t-1}\eta_k + 1}{2}\|\tilde{\boldsymbol{\lambda}}\|_2^2\right) = \frac{1}{2(\sum_{k=0}^{t-1}\eta_k + 1)}\left\|\left[\sum_{k=0}^{t-1}\gamma_k\left(\mathbf{A}\boldsymbol{X}_k - \boldsymbol{b}\right)\right]_+\right\|_2^2$$

■

Setting the optimizing $\tilde{\boldsymbol{\lambda}} \geqslant 0$ into (5.9), taking the expectation of the resulted inequality, and noticing that $\mathbb{E}[S_n(\boldsymbol{x}_*)] = 0$, since $S_n(\boldsymbol{x}_*)$ is a martingale with $\mathbb{E}[S_1(\boldsymbol{x}_*)] = 0$, we obtain the desired statement with $\tilde{C}_2 > 0$ a constant satisfying $\|\boldsymbol{\lambda}_*\|_2^2$

On Trackability Condition: In order that (5.3) is fulfilled at a time τ, it is necessary that:

$$\eta_\tau^2 - (\eta_\tau/2) < 0.$$

Therefore, the requirement (5.3) demands that:

$$\eta_\tau < 1/2.$$

This observation gives the advice to the resources not to be fully progressive in the price determination , i.e. to avoid the parameter $\eta_\tau \approx 1$. By attempting to solve the quadratic inequality (5.3) one can see that a necessary condition on γ_τ s.t. (5.3) holds at time τ is:

$$\gamma_\tau \leqslant \frac{\sqrt{2K}}{C_1}.$$

In this case, (5.3) is equivalent to:

$$\frac{\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{8\gamma_\tau^2 C_1^2}{K}}}{2} \leqslant \eta_\tau \leqslant \frac{\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{8\gamma_\tau^2 C_1^2}{K}}}{2}.$$

This observation assert that for small γ_τ, one can choose η_τ approximately in the interval $(0, 1/2)$.

Remark 18: Suppose that $\gamma_\tau = C_\gamma\tau^{-p}$ for a certain C_γ and $p > 0$, and $\eta_\tau = C_\eta\tau^{-q}$ for a certain $C_\eta > 0$ and $q > 0$. In order that TC holds, it is necessary that η_τ decays with the same order like or slower than γ_τ^2. Therefore we have to require $q \in (0, 2p]$.

Now we are ready to give several consequences of Theorem (5.1). For simplicity, we assume that the noise is persistent, in the sense that $\mathbb{E}[\|\boldsymbol{M}_\tau\|_*^2] \leqslant \sigma_*^2$, for all $\tau \in \mathbb{N}$.

Constant Learning rate: Let us consider a finite time horizon $T \in \mathbb{N}$ and $\beta_\tau = 2$, for all $\tau \in [T-1]_0$. Furthermore, let us consider the case where both, the learning rate of the agents and the price progressivity are constant, i.e.:

$$\gamma_\tau = \gamma \quad \text{and} \quad \eta_\tau = \eta, \quad \forall \tau \in [T-1]_0.$$

Assuming that γ and η fulfills (5.3) holds, it follows from (5.4) and Jensen's inequality:

$$\mathbb{E}\left[\|[\sum_{\tau=0}^{t-1}\gamma_\tau(\mathbf{A}\boldsymbol{X}_\tau - \boldsymbol{b})]_+\|_2^2\right] \leqslant 2\overline{\eta}_t\left(\Delta\psi + \tilde{C}_1\sum_{\tau=0}^{t-1}\gamma_\tau^2\right) + \overline{\eta}_t^2\left(\|\boldsymbol{\lambda}_*\|_2^2 + \frac{4}{K}\sum_{\tau=1}^{t}\gamma_{\tau-1}^2\mathbb{E}[\|\boldsymbol{M}_\tau\|_*^2]\right), \tag{5.10}$$

$$\mathbb{E}[\mathrm{ANCCVC}_T] \leqslant \sqrt{\frac{2(\eta T+1)}{\gamma T^2}\left(\frac{\Delta\psi}{\gamma} + \tilde{C}_1\gamma T\right) + \frac{(\eta T+1)^2}{\gamma^2 T^2}\tilde{C}_2^2 + \frac{(\eta T+1)^2}{T^2}\frac{4\sigma_*^2 T}{K}}. \tag{5.11}$$

So, suppose that $\gamma = \Theta(T^{-p})$ with $p \in [1/2, 1)$. Setting $\eta = \Theta(T^{-q})$ where $q \in (1/2, 2p]$, it yields:

$$\mathbb{E}[\mathrm{ANCCVC}_T] \leqslant \mathcal{O}\left(T^{p-\frac{q+1}{2}} + T^{p-q} + \sigma_* T^{\frac{1}{2}-q}\right). \tag{5.12}$$

In particular if $p = 1/2$, we can choose $q = 1$ and obtain a sub-linear bound for the ANCCVC at time T of order $\mathcal{O}((1+\sigma_*)T^{-\frac{1}{2}})$.

Variable Parameters: If the agents are each willing to apply the ergodic average of their historical strategies instead of their actual strategies, we can ensure the decay of the violation of resource constraints with time in expectation. In order to show this, let us consider the infinite time horizon $T = \infty$. We use Jensen's inequality to obtain the following bound from (5.4):

$$\mathbb{E}\left[\|[\sum_{\tau=0}^{t-1}\gamma_\tau(\mathbf{A}\boldsymbol{X}_\tau - \boldsymbol{b})]_+\|_2\right] \leqslant \sqrt{2\overline{\eta}_t\left(\Delta\psi + \tilde{C}_1\sum_{\tau=0}^{t-1}\gamma_\tau^2\right) + \overline{\eta}_t^2\left(\tilde{C}_2^2 + \frac{4\sigma_*^2}{K}\sum_{\tau=0}^{t-1}\gamma_\tau^2\right)}, \quad \forall t \in \mathbb{N}$$

For the ergodic average $\overline{\boldsymbol{X}}_t^\gamma = \frac{\sum_{\tau=0}^{t-1}\gamma_\tau \boldsymbol{X}_\tau}{\sum_{\tau=0}^{t-1}\gamma_\tau}$ of the population iterate, we have:

$$\mathbb{E}\left[\|[\mathbf{A}\overline{\boldsymbol{X}}_t^\gamma - \boldsymbol{b}]_+\|_2\right] \leqslant \frac{\sqrt{2\overline{\eta}_t\left(\Delta\psi + \tilde{C}_1\sum_{\tau=0}^{t-1}\gamma_\tau^2\right) + \overline{\eta}_t^2\left(\tilde{C}_2^2 + \frac{4\sigma_*^2}{K}\sum_{\tau=0}^{t-1}\gamma_\tau^2\right)}}{\sum_{\tau=0}^{t-1}\gamma_t},$$

Setting $\gamma_t = \Theta(t^{-1/2})$ and $\eta_t = \Theta(t^{-1})$ fulfilling trackability condition, it follows that the decay of the congestion state is in the noiseless case of order $\mathcal{O}(\ln(t)/\sqrt{t})$ and otherwise

$\mathcal{O}(\ln^{3/2}(t)/\sqrt{t})$. Now let be $\gamma_t = \Theta(t^{-1})$ and $\eta_t = \Theta(t^{-2})$, we have decay of order $\mathcal{O}((1 + \sigma_*)/\ln(t))$.

5.4. Numerical Experiment

Exponential Weights Online Learning in Quadratic Game: We consider N agents whose task is to allocate a certain amount of tasks to R resources. The strategy space of agent i corresponds to the simplex $\Delta := \left\{x^{(i)} \in \mathbb{R}^R_{\geqslant 0} : \sum_{r=1}^R \boldsymbol{x}_k^{(i)} = 1\right\}$. For a strategy $x^{(i)} \in \Delta$, $\boldsymbol{x}_r^{(i)}$ stands for the proportion of tasks agent i assigns to resource $r \in [R]$. The cost function of player i is quadratic and given by $J^{(i)}(x^{(i)}, x^{(-i)}) = \frac{1}{2}\langle x^{(i)}, Qx^{(i)}\rangle + \langle C\sigma(\boldsymbol{x}) + c^i, x^{(i)}\rangle$, where $\sigma(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^N x^{(i)}$

where $c_i \in \mathbb{R}^D$, $Q \in \mathbb{R}^{D\times D}$ and $C \in \mathbb{R}^{D\times D}$ are positive semi-definite, and either Q or C are positive definite. In order to apply our method, we set $u^{(i)}(\boldsymbol{x}) = -J^{(i)}(\boldsymbol{x})$. The corresponding gradient mapping is given by

$$\mathbf{v}(\boldsymbol{x}) = -\left[(I_N \otimes Q + \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^{\mathrm{T}} \otimes C)x + c + \frac{1}{N}(I_N \otimes C^{\mathrm{T}})x\right],$$

where $\otimes$ denotes the Kronecker product between two matrices. For the mirror map of the agents, we use the logit choice implemented by means of log-sum trick in order to avoid numerical overflow.

Game Parameter: We consider $N = 20$, $D = R = 5$, and $T = 500$, and study the case where parameters are fixed. We set $Q = 2\sqrt{\tilde{Q}^{\mathrm{T}}\tilde{Q}} + \mathbf{I}_D$, where the entries of $\tilde{Q}$ is chosen independently normal distributed. Moreover we consider the case where $C = 4\mathbf{I}_D$, $c = 0$. For specific model of the stochastic feedback we use Gaussian vector with covariance matrix $\sigma^2\mathbf{I}_D$, where $\sigma > 0$.

Evaluation:

Figure 5.2 shows that pure egoistic uncontrolled behaviour of the agents ($\beta = 0$) may lead to immense overuse of the resources, and that control of agents' consumption via price mechanism ($\beta > 0$) can prevent this event. With price regularization ($\beta > 0$), we observe the tendency of oscillation in the agents' dynamic, whereby the following difference is observable: The choices $\beta = 1$ and $\beta = 2$ effect in stabilizing behaviour, while $\beta = 3$ and $\beta = 4$ effect in chaotic behaviour. This observations are aligned with the gedankenexperiment done in Section 5.2, one of whose conclusions is that high price sensitivity might cause the agents' utilization strategies to mutually bouncing between single resources. Furthermore, Figure 5.2 confirms the optimality of the selection of parameter choice $\beta = 2$ given in Theorem 5.1, since it tends to have the lowest ANCCVC. From Figure 5.1, we can observe that in the non-progressive case $\alpha = 0$ ($\eta = \alpha\gamma^2 = 0$), the corresponding dynamic of ANCCVC resonates heavily and possess at the end of the time-horizon ($t = 500$) highest value (aside from $\alpha = 50$). This asserts the importance

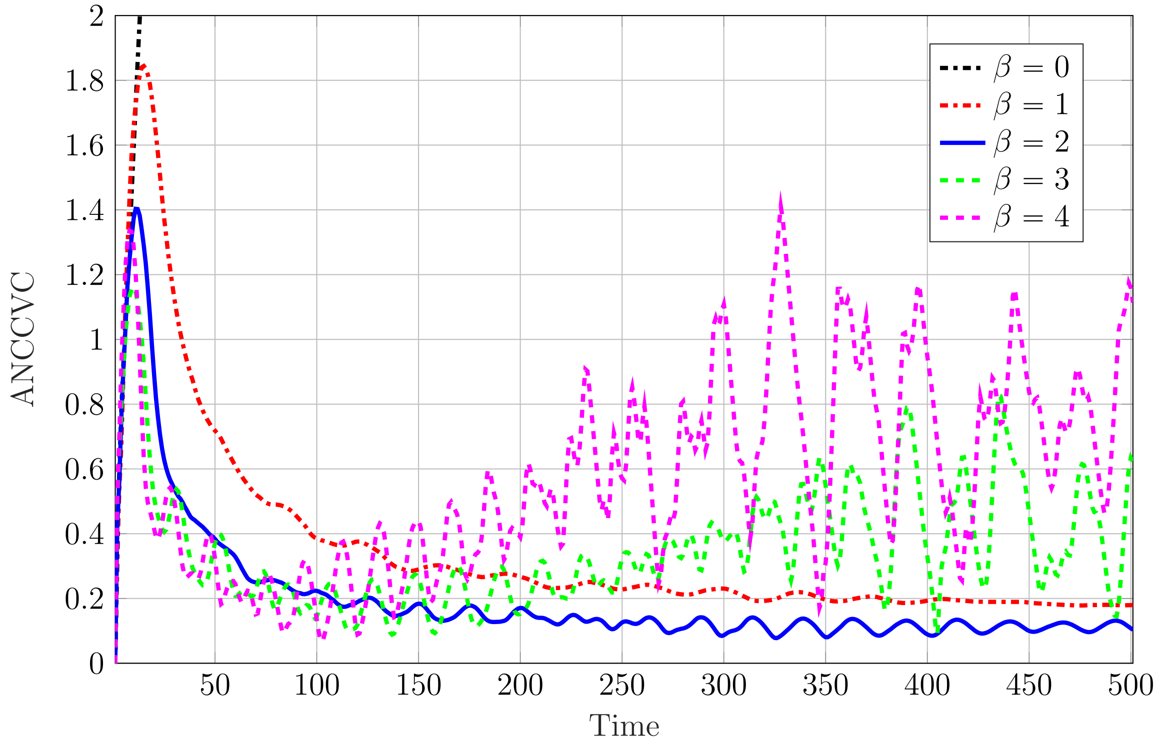


Figure 5.1.: Dynamic of ANCCVC of Algorithm 4 for different price sensitivities β with $\gamma = 0.5/\sqrt{T}$, $\alpha = 10$, and $\sigma = 5$

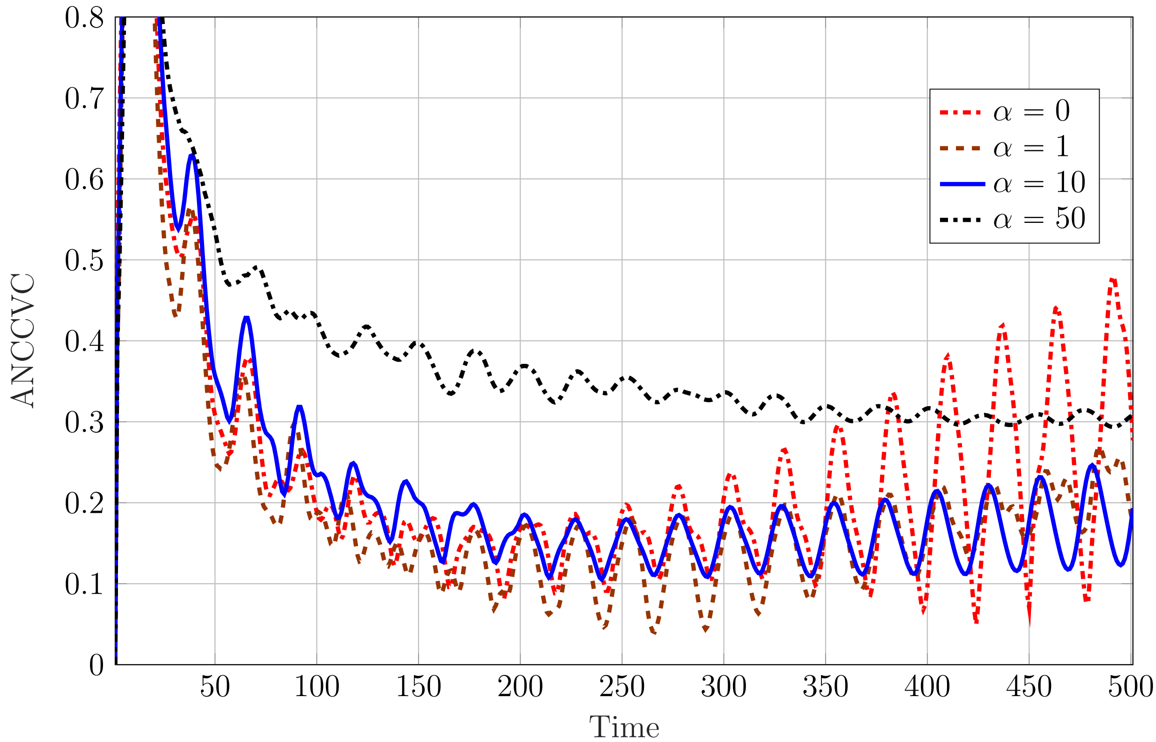


Figure 5.2.: Dynamic of ANCCVC of Algorithm 4 for different price progressivities $\eta = \alpha\gamma^2$ with $\beta = 2$, $\gamma = 0.5/\sqrt{T}$, and $\sigma = 5$

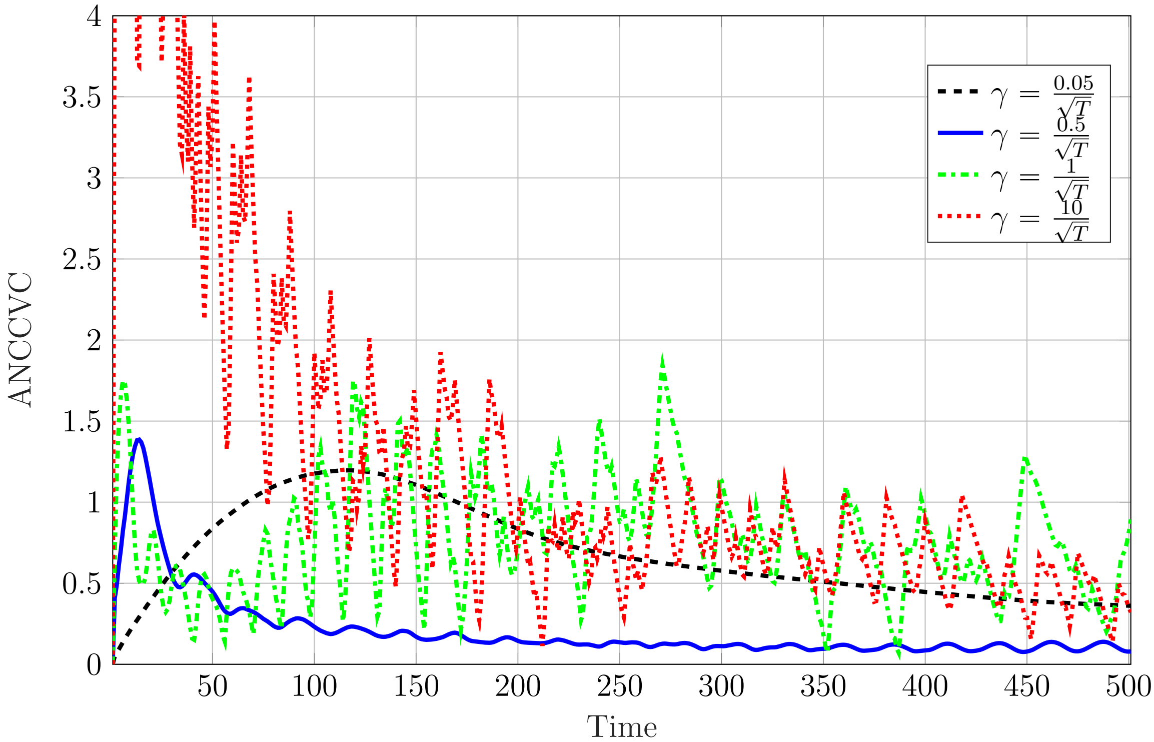


Figure 5.3.: Dynamic of ANCCVC of Algorithm 4 for different learning rates γ with $\beta = 2$, $\alpha = 10$, and $\sigma = 5$

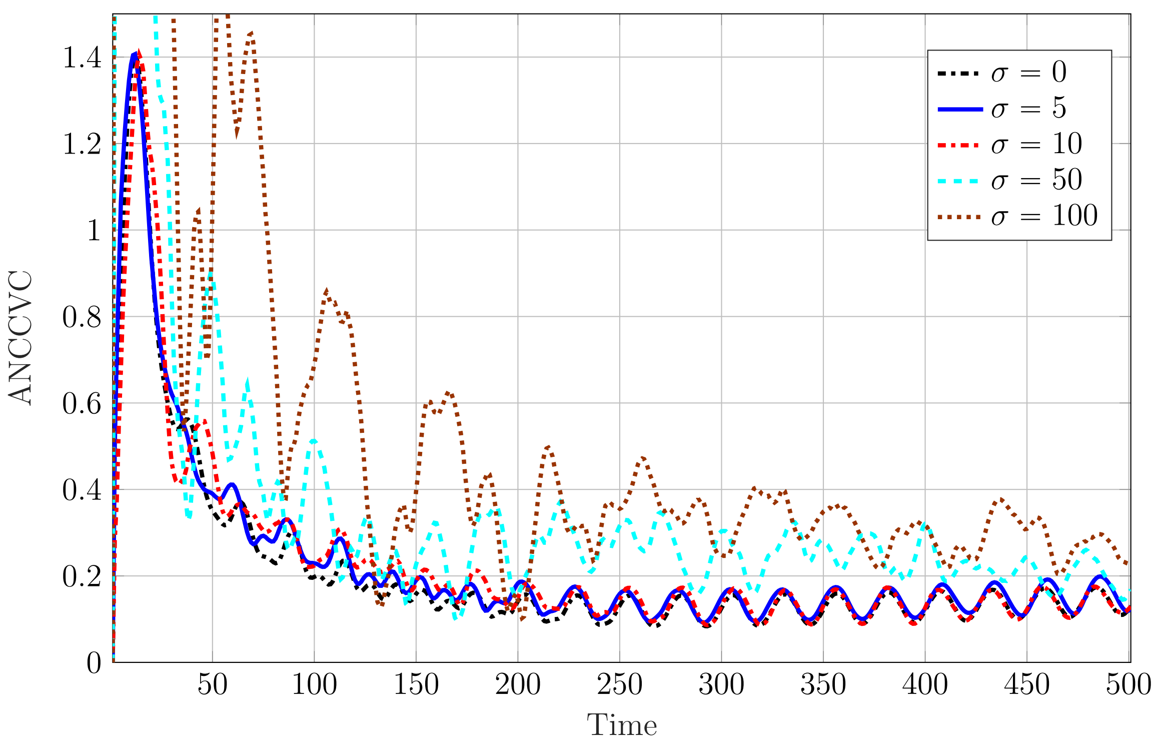


Figure 5.4.: Dynamic of ANCCVC of Algorithm 4 for different noise powers σ with $\beta = 2$, $\alpha = 10$, and $\gamma = 0.5/\sqrt{T}$

of progressivity in the price determination and also justifies the importance of the TC (5.3). The parameters $\alpha = 1$ and $\alpha = 10$ has the best behaviour in this experiment. We see the tendency of decreasing oscillation with increasing price progressivity. However by observing $\alpha = 10$ the overall performance at the end of the time horizon might be worse for high η. This observation underlines the role of η as the parameter specifying the decay rate of ANCCVC asserted by the bound (5.11). From Figure 5.3 the high oscillatory chaotic behaviour of $\gamma = 1/\sqrt{T}$ and $\gamma = 10/\sqrt{T}$ is aligned with the TC (5.3) which eliminate the possibility that for fixed η, γ can be arbitrarily high. Moreover the plot for $\gamma = 0.05/\sqrt{T}$ shows that too small γ caused slow decay of the ANCCVC as predicted by the bound (5.11).

Figure 5.4 shows that the noise power has no significant influence to the ANCCVC. This observation is somehow forecasted by our theoretical results since the noise term in the corresponding expectation bound decay with square roots of the time and the noise is light-tailed.

6. Resource-Aware Control via Pricing for Congestion Game with Finite-Time Guarantees

Abstract: Congestion game is a widely used model for modern networked applications. A central issue in such applications is that the selfish behavior of the participants may result in resource overloading and negative externalities for the system participants. In this work, we propose a pricing mechanism that guarantees the sub-linear cumulative violation of the resource load constraints, of square root order w.r.t. the time horizon. The feature of our mechanism is that it is resource-centric in the sense that it depends on the congestion state of the resources and not on specific states and characteristics of the system participants. This feature makes our mechanism scalable, flexible, and privacy-preserving. Moreover, we show by numerical simulations that our pricing mechanism has no significant effect on the agents' welfare and may even result in the improvement of the latter, depending on the parameter choice.

6.1. Introduction

Modern networked systems such as IoT, smart grid, and cognitive radio are characterized by optimizing users/devices which dynamically compete for utilization of resources, be it network link, power supply, and wireless spectrum. A common trend in recent years is that the number of users in such applications increases tremendously (see e.g. [119]). For instance: Analysts predicts that more than 50 billion things are expected to be connected over the internet by the end of 2020 [119]. Such rapid growth involves certainly a series of challenges.

One of the main challenges facing the system managers is the congestion control of the available resources. For without it, negative externalities in the form of immense degradation of the quality of service of the resources might occur due to overload. For instance, in wireless communication network applications, if the amount of traffic through a router (resource) exceeds its capacity, buffer bloat occurs, resulting in inefficiency of the system in the form of high latency and network throughput reduction, causing a negative experience for all users. Moreover, sophisticated congestion control method is crucial for making the electrical power driven technologies environment-friendly, Another trend visible in recent years is that power consumption due to technical applications constitutes a non-negligible part of the global power consumption with the tendency of enormous growth (see e.g., [120]).

The concept of the congestion game introduced in [121, 122] is a natural fundament for developing a congestion control method. The corresponding model assumes non-cooperative rational participants, whose strategy is an allocation policy over a collection of the subset of resources and whose loss depends proportionally on the total load of the utilized resources. The most prominent classical example of a congestion game is the traffic routing model of Wardrop [123], where the arcs in a given network represent the resources, the different origin-destination pairs specify the player, and the possible action of a player is the allocation over the paths in the network between his origin-destination pair.

Many congestion control methods are user-centric in the sense that they require observability of system participants' actions and behaviors and provide specific instructions for all of the system users. Such methods are not suitable for modern large-scale applications. The reason is threefold: First, such methods often lack scalability and flexibility. Second, the typically high number of participants in such applications makes the methods computationally infeasible. Third, due to growing users' demands of sovereignty and privacy in recent years, direct observation and influence of users' acts by higher authority are highly undesirable.

Our Contributions Based on the assumption of rational non-cooperative cost-oriented agents, we propose resource-centric dynamic pricing that offers the system participants appropriate incentives to adhere to the resource constraints jointly support sustainable use of the resources. In this context, resource-centric means that the given method is mainly based on the observation of the actual congestion of the resource and not on the agents' specific characteristics such as the set of their possible bundle choices and their actual bundle choices. We show that the proposed pricing mechanism ensures that the average violation of the capacity constraints decays at worst sub-linearly of order $\mathcal{O}(n^{-1/2})$ w.r.t. the time-horizon n. Moreover, we provide numerical simulations in order to support our theoretical findings. As a by-product of our practical investigations, we observe that, although it does not use specific information about the agents, our pricing mechanism does not effectuate the agents' welfare, expressed by their average loss, significantly, and may even result in improvement of the latter.

Relation to prior work The congestion game without resource constraints has been investigated in several directions. Towards this aspect, the following approaches, which consider the game to be played multiple times, are closely related to our work: Under different black-box behavior of the individual agents, [124–126] study the convergence of selfish choice toward the Nash equilibrium.

One can trace back the idea of mitigating efficiency due to negative externalities by pricing method to [1], whose main advice is to charge externalities-causing agents additional costs and to reallocate the obtained compensation to the agents sustaining the externalities in order that the population establishes an optimal social state. In contrast to our work, the Pigouvian method requires that the price maker knows about the preferences of every agent to determine the socially optimal state, which is infeasible in large scale technical applications.

Several works introduce exciting approaches to game-theoretic pricing based congestion control methods (to name a few: [68, 84, 127–129]. However, in contrast to our work, they only provide an asymptotic guarantee by designing a population dynamic which converges to the corresponding (designed) equilibrium fulfilling the capacity constraints (see e.g., the concept of generalized Nash equilibrium [93]) of the problem-specific potential game [130]. Until now, there is no approach to design a congestion control method that possesses a guarantee in the non-asymptotic regime. Specifically they are only able to guarantee the preservation of the constraint in large time. In the literature, the cost function of the agents consists, in contrast to our modeling, not only of congestion loss - and price-dependent term but also of idiosyncratic payoff, which depends only on the action of a single agent. However the given analysis in this paper can be extended straightforwardly to this model.

6.2. Setting

6.2.1. Congestion Game with Resource Constraints

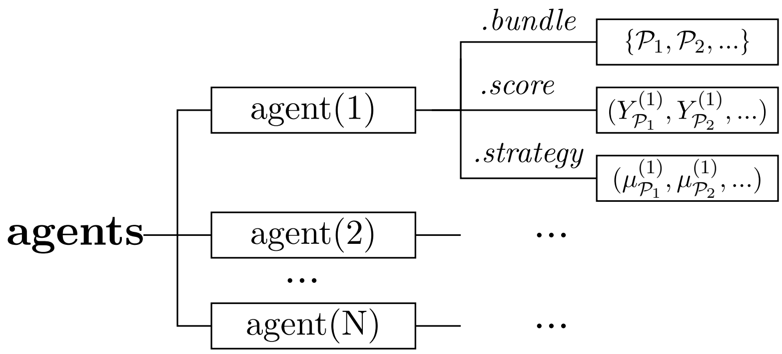


Figure 6.1.: Agent Structure

A congestion game consists of a finite set of agents/players $[N]$ and a finite set $\mathcal{R}$ of resources. To each agent $i \in [N]$, there corresponds a collection $\mathfrak{P}_i \subseteq 2^{\mathcal{R}}$ of resource bundles. Her aim is to execute a certain amount $m_i > 0$ of tasks by utilizing the bundles of resources from $\mathfrak{P}_i$. We describe the corresponding (utilization) action/strategy of agent

i by a vector $\boldsymbol{x}^{(i)} \in \mathcal{X}_i$, where $\mathcal{X}_i$ is a scaled simplex on $\mathfrak{P}_i$, i.e.:

$$\mathcal{X}_i := \left\{ \boldsymbol{x}^{(i)} := (\boldsymbol{x}^{(i)}_{\mathcal{P}_i})_{\mathcal{P}_i \in \mathfrak{P}_i} \in \mathbb{R}^{|\mathfrak{P}_i|} : \sum_{\mathcal{P}_i \in \mathfrak{P}_i} \boldsymbol{x}^{(i)}_{\mathcal{P}_i} = m_i \right\}.$$

For any $\mathcal{P}_i \in \mathfrak{P}_i$, $\boldsymbol{x}^{(i)}_{\mathcal{P}_i}$ corresponds to the amount of tasks agent i allocates to the bundle $\mathcal{P}_i$. Equivalently, we can describe the task allocation strategy of agent i by means of the simplex

$$\Delta_i := \left\{ \boldsymbol{\mu}^{(i)} := (\boldsymbol{\mu}^{(i)}_{\mathcal{P}_i})_{\mathcal{P}_i \in \mathfrak{P}_i} \in \mathbb{R}^{|\mathfrak{P}_i|} : \sum_{\mathcal{P}_i \in \mathfrak{P}_i} \boldsymbol{\mu}^{(i)}_{\mathcal{P}_i} = 1 \right\}.$$

In this paper we describe the allocation strategy of agent i by means of the simplex Δ_i instead with $\mathcal{X}_i$. We denote the set of population strategy by:

$$\Delta = \prod_{i=1}^{N} \Delta_i.$$

Let $\boldsymbol{\mu}^{(i)} \in \Delta_i$ be an allocation action of agent i. The total load $\boldsymbol{\phi}^{(i)}_r(\boldsymbol{\mu}^{(i)})$ of the resource $r \in \mathcal{R}$ caused by the allocation action $\boldsymbol{\mu}^{(i)} \in \Delta_i$ of agent i is given by:

$$\boldsymbol{\phi}^{(i)}_r(\boldsymbol{\mu}^{(i)}) = \sum_{\mathcal{P}_i \in \mathfrak{P}_i : r \in \mathcal{P}_i} m_i \boldsymbol{\mu}^{(i)}_{\mathcal{P}_i} = (\widetilde{\mathcal{M}}^{(i)} \boldsymbol{\mu}^{(i)})_r,$$

where:

$$\widetilde{\mathcal{M}}^{(i)} = m_i \mathcal{M}^{(i)} \quad \text{and} \quad \mathcal{M}^{(i)} \in \mathbb{R}^{|\mathcal{R}| \times |\mathfrak{P}_i|}$$

is the adjacency matrix whose $\mathcal{P}_i$-th column provides the information about all the resources contained in the bundle $\mathcal{P}_i$, i.e.:

$$[\mathcal{M}^{(i)}]_{r,\mathcal{P}_i} = \begin{cases} 1 & r \in \mathcal{P}_i \\ 0 & \text{else} \end{cases}.$$

Accordingly, the total load $\boldsymbol{\phi}_r(x)$ of resource r caused by the population strategy $\boldsymbol{\mu} \in \Delta$ is given by:

$$\boldsymbol{\phi}_r(\boldsymbol{\mu}) = \sum_{i=1}^{N} \boldsymbol{\phi}^{(i)}_r(\boldsymbol{\mu}^{(i)}). \tag{6.1}$$

We sometimes also use the notation $\boldsymbol{\phi} := (\boldsymbol{\phi}_r)_{r \in \mathcal{R}}$. One can express (6.1) more compactly by

$$\boldsymbol{\phi}(\boldsymbol{\mu}) = \widetilde{\mathcal{M}} \boldsymbol{\mu}, \quad \text{where} \quad \widetilde{\mathcal{M}} = [\widetilde{\mathcal{M}}^{(1)} | \cdots | \widetilde{\mathcal{M}}^{(N)}].$$

Additionally we consider the case where the load of resources $r \in \mathcal{R}$ is desirable to not exceed the capacity $L_r \in \mathbb{R}_{>0}$, i.e.:

$$\boldsymbol{\phi}_r(\boldsymbol{\mu}(k)) - L_r =: \boldsymbol{\Gamma}_r(\boldsymbol{\mu}) \leqslant 0$$

Specifically, it is desired that the population strategy $\boldsymbol{\mu} \in \Delta$ satisfies $\boldsymbol{\mu} \in \mathcal{Q}$, where:

$$\mathcal{Q} := \{\boldsymbol{\mu} \in \Delta : \ \boldsymbol{\Gamma}(\boldsymbol{\mu}) \leqslant 0\},$$

with $\boldsymbol{\Gamma}(\boldsymbol{\mu}) := (\boldsymbol{\Gamma}_r(\boldsymbol{\mu}))_{r\in\mathcal{R}}$ a vector with positive entries.

To each resource $r \in \mathcal{R}$, we associate a function $\ell_r : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ which quantifies negative externalities induced on the resource r due to load $\boldsymbol{\phi}_r(\boldsymbol{\mu})$. We refer to ℓ_r as the loss function of the resource r. We assume throughout that the following properties holds:

Assumption 6.1: *For all $r \in \mathcal{R}$, $\ell_r : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ is continuous, convex, and non-decreasing.*

Assumption 6.2 (Slater's Condition): *There exists $\hat{\boldsymbol{\mu}} \in \Delta$ s.t. $\boldsymbol{\Gamma}(\hat{\boldsymbol{\mu}}) < 0$.*

The loss of a bundle $\mathcal{P}_i \in \mathfrak{P}_i$ (for agent i) is correspondingly given by:

$$\ell^{(i)}_{\mathcal{P}_i}(\boldsymbol{\mu}) = \sum_{r\in\mathcal{P}_i} \ell_r(\boldsymbol{\phi}_r(\boldsymbol{\mu})),$$

Throughout this work we use the notations $\ell^{(i)} := (\ell^{(i)}_{\mathcal{P}_i})_{\mathcal{P}_i\in\mathfrak{P}_i}$ and $\ell := (\ell^{(i)})_{i\in[N]}$.

An example of congestion game is the following:

Example 12 (Network Routing Game): Given a directed Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a vertex set $\mathcal{V}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. In a routing game, the task of agent $i \in [N]$ is to transport a certain amount of commodity $m_i > 0$ from a starting point $s^{(i)} \in \mathcal{V}$ to a destination $t^{(i)} \in \mathcal{V}$. To fulfill this task, agent i can use a prescribed collection $\mathfrak{P}_i \subseteq 2^{\mathcal{E}}$ of edges that connects $s^{(i)}$ and $t^{(i)}$. To every edge (resource) $e \in \mathcal{E}$ there corresponds a function c_e (loss) that maps the total amount flow caused by the transport of commodities on the edge e to a non-negative number determining the delay on e, and also a constant $\ell_e > 0$ which prescribed the amount of flow admissible on edge e.

6.2.2. Performance Measures

Let be $k \in \mathbb{N}$ and $\boldsymbol{\mu}(\tau)$, $\tau \in [k]_0$, be a given sequence of population actions from initial time until time slot k. To evaluate the population performance in the congestion game we use the following criteria:

We measure the resource sustainability of the population sequential actions $(\boldsymbol{\mu}(\tau))_{\tau\in[k]_0}$ by the (norm) of the aggregated admissible flow violation defined by:

$$\mathrm{ACV}(k) = \left\| \left[\sum_{\tau=0}^{k-1} \boldsymbol{\Gamma}_r(\boldsymbol{\mu}(\tau)) \right]_+ \right\|_2$$

Additional to resource sustainability behavior, we investigate the loss incurred to the population applying the resource allocation decisions $(\boldsymbol{\mu}(\tau))_{\tau\in[k]_0}$ in form of the aggregated delay:

$$\mathrm{AD}(k) = \sum_{\tau=0}^{k} \sum_{i\in[N]} \mathrm{D}_i(\tau),$$

where $\mathrm{D}_i(\tau)$ denotes the delay experienced by agent i at time τ:

$$\mathrm{D}_i(\tau) = \sum_{\mathcal{P}_i\in\mathfrak{P}_i} \ell^{(i)}_{\mathcal{P}_i}(\boldsymbol{\mu}(\tau))\boldsymbol{\mu}^{(i)}_{\mathcal{P}_i}(\tau).$$

It should be noted, that resource sustainability and loss minimization do not need to be coinciding objectives, but can display a trade-off behavior depending on model parameters, i.e. they appear as conflicting objectives. Therefore it can happen, that resource sustainability subsequently implies a disadvantaging of some agents.

6.3. Resource-Centric Pricing for Congestion Game

6.3.1. Population Dynamic via Score and Hedge strategy

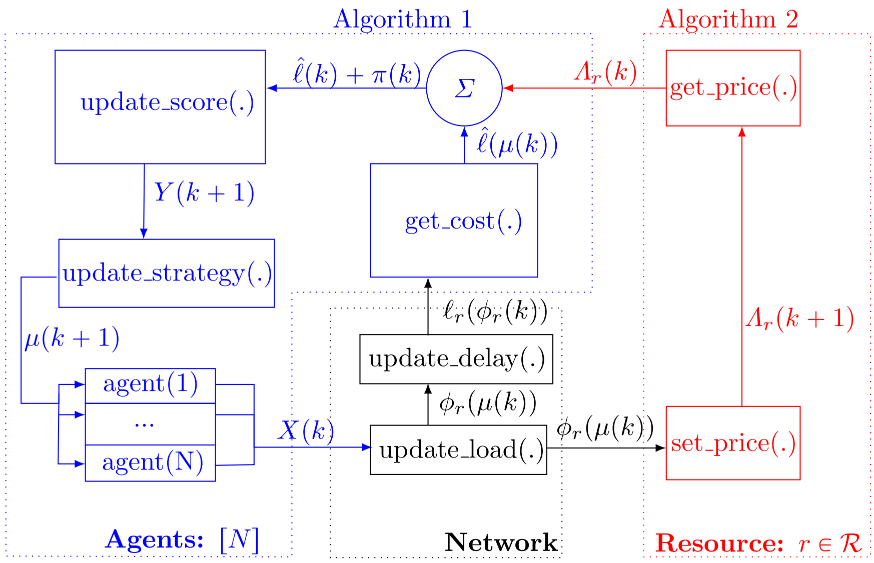


Figure 6.2.: Sketch of Algorithms 1 and 2

Throughout this work, we consider the congestion game, which is played multiply with time horizon $n \in \mathbb{N}$. We give a brief description of our proposal of population dynamic in congestion game in Algorithm 5.

There, at each round $k \in [n]$ every agent $i \in [N]$ accumulates the cost of each resource bundle available to him as to provide its score which reflects his bundle preference. The corresponding actual cost of an available bundle consists of the actual noisy loss caused by negative externalities and the price set exogenously by a regulator.

Algorithm 5

Require: $n \in \mathbb{N}$, $\gamma > 0$, $\mathbf{\Phi}_i : \mathbb{R}^{\mathfrak{P}_i} \to \Delta_i$.

for every agent $i \in [N]$ **do**

Initialize the score vector $Y_0^{(i)} \leftarrow 0$

end for

for time $k = 1, 2, \dots, n$ **do**

Population apply the allocation strategy:

$$\boldsymbol{X}(k) = (m_i \boldsymbol{\mu}^{(i)}(k))_{i \in [N]}$$

for every agent $i \in [N]$ **do**

Receive the price vector $(\boldsymbol{\Lambda}_r(k))_{r \in \mathcal{R}}$ broadcasted by the regulator.

for all bundle of resource $\mathcal{P}_i \in \mathfrak{P}_i$ **do**

Experience the disturbed cost:

$$\hat{\ell}_{\mathcal{P}_i}^{(i)}(k) \leftarrow \ell_{\mathcal{P}_i}^{(i)}(\boldsymbol{\mu}(k)) + \boldsymbol{M}_{\mathcal{P}_i}^{(i)}(k+1)$$

Compute the price per amount of task:

$$\pi_{\mathcal{P}_i}^{(i)}(k) = \sum_{r \in \mathcal{P}_i} \boldsymbol{\Lambda}_r(k)$$

Update the score of bundle $\mathcal{P}_i$:

$$Y_{\mathcal{P}_i}^{(i)}(k+1) \leftarrow Y_{\mathcal{P}_i}^{(i)}(k) - \gamma \left[\hat{\ell}_{\mathcal{P}_i}^{(i)}(k) + \pi_{\mathcal{P}_i}^{(i)}(k) \right] \tag{6.2}$$

end for

Generate the allocation strategy:

$$\boldsymbol{\mu}^{(i)}(k+1) \leftarrow \mathbf{\Phi}^{(i)}(Y^{(i)}(k+1))$$

end for

end for

The noise model we consider is quite general: $(\boldsymbol{M}_n)_{n\in\mathbb{N}}$ is a $\mathbb{R}^{\sum_{i=1}^{N}|\mathfrak{P}_i|}$-valued $\mathbb{F}$-*martingale difference sequence*. One reason that we model the loss as noisy is that the environment or the imperfectness of agents' sensing devices can cause imperfectness of agents' feedback. Another reason is that we can handle the case where the agents' actions are not continuous while their strategies are mixed states and thus the resource congestion only represents an unbiased sample of the congestion specified by the mixed strategies (see [76]).

The parameter γ specifies the step size of the iterates of the agents and quantifies the conservativeness of the agents. The mapping $\boldsymbol{\Phi}^{(i)}$ serves to model how the i^{th} agent builds up his allocation strategy from the actual score of the bundles. In this work, we investigate the case where it takes the following specific form:

$$(\boldsymbol{\Phi}^{(i)}(y^{(i)}))_{\mathcal{P}_i} = \frac{\exp(y^{(i)}_{\mathcal{P}_i})}{\sum_{\hat{\mathcal{P}}_i\in\mathfrak{P}_i}\exp(y^{(i)}_{\hat{\mathcal{P}}_i})}. \tag{6.3}$$

Notice that it holds $\boldsymbol{\Phi}_i > 0$. So every agent utilizes all possible bundles although some might be underutilized. This diversification strategy ensures in particular that each agent keeps track of the congestion state and therefore, the loss of all bundles.

Remark 19: Without altering the analysis given in this work, one can use more generally the concept of mirror map (see Definition 2.1)

6.3.2. Pricing Algorithm

In order to encourage sustainable use of the resources, we proposed the pricing mechanism for Algorithm 5 in Algorithm 6.

Algorithm 6

Require: $n\in\mathbb{N}$, $\beta>0$, $\alpha\in(0,1]$
  Initialize the price vector $\boldsymbol{\Lambda}_0 \leftarrow 0$
  for time $k=1,2,\dots,n$ **do**
    for Regulator **do**
      for $r\in\mathcal{R}$ **do**
        Check the actual load $\boldsymbol{\phi}_{r,k} := \boldsymbol{\phi}_r(X(k))$ of resource r caused by Algorithm 5
        Update the price of resource r:

$$\boldsymbol{\Lambda}_r(k+1) \leftarrow \left[(1-\alpha)\boldsymbol{\Lambda}_r(k) + \beta\left(\boldsymbol{\phi}_{r,k} - L_r\right)\right]_+$$

      end for
    end for
  end for

In case that $\alpha = 1$, then the price update is based entirely on the actual congestion of the resources. One can interpret The updating rule with this choice of the parameter can be interpreted as clinging too much to actual congestion observation. If $\alpha < 1$ then the actual price vector is considered in the price update. This parameter choice reflects the conservativeness of the regulator. Those observations lead us to call α the response parameter.

Another reason to design the price as above is that by that way, one can track the aggregate dynamic of the congestion state by observing the price:

Lemma 6.1: *Suppose that $\boldsymbol{\Lambda}_0 = 0$. For all $r \in [R]$ and $k \in \mathbb{N}$:*

$$\mathrm{ACV}(k) \leqslant \frac{\|\boldsymbol{\Lambda}(k)\|_2 + \alpha \sum_{\tau=1}^{k-1} \|\boldsymbol{\Lambda}(\tau)\|_2}{\beta} \tag{6.4}$$

The proof of this statement can be found in the appendix.

6.4. Performance Analysis

Throughout, C_1, C_2, C_3, m_* denote non-negative constants fulfilling for all $\boldsymbol{\mu} \in \Delta$ and $\boldsymbol{\Lambda} \in \mathbb{R}^{\mathcal{R}}_{\geqslant 0}$:

$$\begin{aligned} &\sum_{i=1}^{N} m_i \|M^{(i),\mathrm{T}} \boldsymbol{\Lambda}\|_\infty^2 \leqslant C_1^2 \|\boldsymbol{\Lambda}\|_2^2, \ \sum_{i=1}^{N} m_i \|\ell^{(i)}(\boldsymbol{\mu})\|_\infty^2 \leqslant C_2^2 \\ &\|\boldsymbol{\phi}(\boldsymbol{\mu}) - L\|_2 \leqslant C_3, \quad m_i \leqslant m_*, \ \forall i \in [N]. \end{aligned} \tag{6.5}$$

Our main result is the following:

Theorem 6.2: *Let $\gamma > 0$ be given, $\beta = \gamma$, and $\alpha = \delta\gamma^2$ with $\delta >$ satisfying*

$$(C_1^2 + \gamma^2 \delta^2) - \frac{\delta}{2} \leqslant 0. \tag{6.6}$$

It holds:

$$\begin{aligned} \mathbb{E}\left[\frac{\|\boldsymbol{\Lambda}(n) - \boldsymbol{\lambda}_*\|_2^2}{2}\right] \leqslant & \frac{\Delta\psi^2}{2} + (1 + \alpha n)\frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} + \frac{\tilde{C}_1^2}{2}\gamma^2 n \\ & + 2\gamma^2 m_* N \sum_{k=1}^{n} \mathbb{E}[\|\boldsymbol{M}_k\|_\infty^2] \end{aligned} \tag{6.7}$$

where $\widetilde{C}_1^2 := 2\left(C_2^2 + 2C_3^2\right)$ and $\Delta\psi^2 = 2m_* \sum_{i=1}^{N} \ln(|\mathfrak{P}_i|)$

The poof of Theorem 6.2 is given in the appendix.

An immediate consequence of above result is the following (for proof see the appendix):

Corollary 6.3: *Suppose that the conditions of Theorem 6.2 are fulfilled and that the noise is persistent in the sense that there exists $\sigma^2 > 0$ s.t. $\mathbb{E}[\|\boldsymbol{M}_k\|_\infty^2] \leqslant \frac{\sigma^2}{4m_* N}$ for all $k \in \mathbb{N}$. It holds:*

$$\begin{aligned} \mathbb{E}\left[\|\boldsymbol{\Lambda}(n)\|_2\right] \leqslant & \Delta\psi + (1 + \sqrt{(1 + \delta\gamma^2 n)})\|\boldsymbol{\lambda}_*\|_2 \\ & + (\tilde{C}_1 + \sigma)\gamma\sqrt{n}, \end{aligned} \tag{6.8}$$

where $\Delta\psi$ and $\tilde{C}_1$ is given as in Theorem 6.2. Now, suppose that $\gamma := c/\sqrt{n}$: for a constant $c > 0$ and $\delta \in (0, 1/\gamma^2)$ s.t. (6.6) is fulfilled. It holds:

$$\mathbb{E}\left[\mathrm{ACV}(n)\right] \leqslant (\delta c + \frac{1}{c})A\sqrt{n}, \tag{6.9}$$

where $A := \Delta\psi + (1 + \sqrt{(1 + \delta c^2)})\|\boldsymbol{\lambda}_*\|_2 + (\tilde{C}_1 + \sigma)c$.

6.5. Simulation

Game Setting: We consider the network routing problem given in Example 12 which we specifies as follows: $\mathcal{V}$ consists of 15 nodes and $\mathcal{E}$ is built from a randomly generated adjacency matrix (without self-loop) with independent entries, where each non-diagonal is 1 with probability 0.5. Furthermore, we consider $N = 10$ agents, each has the starting point and the destination randomly uniformly chosen from $\mathcal{V}$. Given the latter, each agent i has randomly created bundles of maximal size $|\mathcal{P}_i| \leqslant 10$. We set the total resource load $m_i = 20$, $\forall i \in [N]$, and the admissible flow per resource $L_r = 14$, $\forall r \in \mathcal{R}$. For the cost per resource $\ell_{r,k}$, we consider a quadratic polynomial of the form $\ell_{r,k}(\boldsymbol{\phi}_r(k)) = a_2^{(r)}\boldsymbol{\phi}_r(k)^2 + a_1^{(r)}\boldsymbol{\phi}_r(k) + a_0^{(r)}$, where the coefficients $(a_2^{(r)}, a_1^{(r)}, a_0^{(r)})$ for each resource $r \in \mathcal{R}$ are independently randomly uniformly chosen from $[0, 0.1]$.

Parameter Setting: We set the parameters required by Algorithms 1 and 2 as follows: We consider the time horizon $n = 10^3$, the agents' learning rate $\gamma = 0.1\sqrt{n} = 0.0032$, and the response parameter $\alpha = 10^{-5}$. We are not only interested in the case $\beta = \gamma$ analyzed in Section 6.4, but also in the case where the regulator is uncertain about the agents' learning rate, and therefore β differs significantly from γ by the factor 10: $\beta = 10\gamma$ ($\beta > \gamma$) and $\beta = 10^{-1}\gamma$ ($\beta < \gamma$).

Performance Evaluation: Fig. 6.3 (b) and (c) show that our pricing mechanism reduces the aggregated capacity violation even if $\beta \neq \gamma$. However, we observe that a higher β may accelerate this process. Additionally, we see that our pricing method does not yield significant discrimination of the agents, as the differences between the aggregated delays for the different cases are marginal at worst (see Fig. 6.3 (a)). For the theoretically analyzed case $\beta = \gamma$, the agents experience, on average even less delay compared to the case of no pricing. Furthermore, we note a trade-off behavior in the choice of β: In case that β is high ($\beta > \gamma$), the capacity violation is the lowest, but the experienced delay the highest. This occurrence reflects the increasing dominance of the price regulation over the agents' personal interest to decrease the incurred delay. Another observation which we make is that if $\beta = \gamma$, some prices might at worst be constant for large times as predicted in Corollary 6.3, indicating that even the population fulfills resource constraints, a control mechanism is necessary to maintain this desired status quo.

Overly Strict Capacity Constraints: We also investigate the performance of our method with stricter capacity constraints than before ($L_r = 11$). We see that our method yields still an improvement of the capacity violation compared to the no pricing case (see Fig. 6.5 (b) and (c)). However, this comes with a significant reduction of agents' welfare in the form of a higher AD (see Fig. 6.5 (a)). One may justify this effect as follows: Taking a look at the pricing evolution (Fig. 6.4 (b)) of exemplary resources, we observe a linear increase in prices dominating the personal preferences ($\hat{\ell}^{(i)}_{\mathcal{P}_i}$ in (6.2)) of the agents in large times. Consequently, each of the affected agents decides for routes that have the lower prices rather than those that incur her the lowest delay.

The enormous increase of prices shown in Fig. 6.5 (a) gives a hint that the minimizer of the Rosenthal potential corresponding to the network routing game over $\mathcal{Q}$ does not exists (c.f. the Proof of Theorem (6.2)) due to overly strict resource constraints. However, one may able to show the sub-linearity of ACV however of order $\mathcal{O}(n^{1/4})$.

Moreover, The increase in prices is in contrast to the case where the capacity constraints are rather loose (Fig. 6.4 (b)). The latter observations give the following heuristic: In case that one observes a linear increase of some prices, one may set a looser constraint so that the reduction of the capacity violations does not come with a significant reduction of the populations' welfare.

6.6. Appendix

6.6.1. Proof of the main result

Proof (Proof of Theorem 6.2): The logit choice $\mathbf{\Phi}^{(i)}$ given in (6.3) is a mirror map (Definition 3.1 in [76]) induced by the negative Gibbs entropy $\psi_i(\boldsymbol{\mu}^{(i)}) = \sum_{\mathcal{P}_i \in \mathfrak{P}_i} \boldsymbol{\mu}^{(i)}_{\mathcal{P}_i} \ln(\boldsymbol{\mu}^{(i)}_{\mathcal{P}_i})$ as regularizer on the simplex which is a compact convex subset. Let be $F^{\mathbf{m}}(\boldsymbol{\mu}, \boldsymbol{Y}(k)) := \sum_{i=1}^{N} m_i F_i(\boldsymbol{\mu}^{(i)}, Y^{(i)}(k))$ where F_i is the Fenchel coupling (Definition 4.2 in [76]) induced by the negative Gibbs entropy as 1-strongly (w.r.t. $\|\cdot\|_\infty$) convex regularizer on the simplex Δ_i.

By means of $F^{\mathbf{m}}$, we can estimate the evolution of Algorithm 5 with the dynamic pricing mechanism given in Algorithm 6 by means of Lyapunov's type argumentation. Toward this end, we use the usual bound for the one step difference of the Fenchel coupling (see e.g. Proposition 4.3 (c) in [76]), insert the given iterate at time $k+1$ in the resulted inequality, and apply triangle inequality, to obtain:

$$\begin{aligned} F^{\mathbf{m}}(\boldsymbol{\mu}, \boldsymbol{Y}(k+1)) - F^{\mathbf{m}}(\boldsymbol{\mu}, \boldsymbol{Y}(k)) \leqslant & -\gamma \underbrace{\sum_{i=1}^{N} m_i \langle \boldsymbol{\mu}^{(i)}(k) - \boldsymbol{\mu}^{(i)}, \hat{\ell}^{(i)}(k) + \pi^{(i)}(k) \rangle}_{=:(\mathrm{a})} \\ & + \frac{\gamma^2}{2} \underbrace{\sum_{i=1}^{N} m_i \|\hat{\ell}^{(i)}(k) + \pi^{(i)}(k)\|_\infty^2}_{=:(\mathrm{b})} . \end{aligned} \tag{6.10}$$

By the triangle inequality and the definition of the constants given in Section 6.4, we can estimate the summand (b) as follows:

$$(\mathrm{b})/2 \leqslant C_1^2 \|\boldsymbol{\Lambda}(k)\|_2^2 + 2(C_2^2 + \sum_{i=1}^{N} m_i \|\boldsymbol{M}_{k+1}^{(i)}\|_\infty^2) \tag{6.11}$$

Now to estimate the summand (a), notice that we can write:

$$\sum_{i=1}^{N} m_i \langle \boldsymbol{\mu}^{(i)}(k) - \boldsymbol{\mu}^{(i)}, \pi^{(i)}(k) \rangle = \langle \boldsymbol{\mu}(k) - \boldsymbol{\mu}, \widetilde{\mathcal{M}}^T \boldsymbol{\Lambda}(k) \rangle. \tag{6.12}$$

Combining all the previous observations, we have by summing the resulting inequality over all $k = 0, \ldots, n-1$, and by subsequent telescoping, we obtain an upper bound for the cumulative difference $\mathcal{V}_n^{(1)}(\boldsymbol{\mu}) := F^{\mathbf{m}}(\boldsymbol{\mu}, \boldsymbol{Y}(n)) - F^{\mathbf{m}}(\boldsymbol{\mu}, \boldsymbol{Y}(0))$:

$$\begin{aligned} \mathcal{V}_n^{(1)}(\boldsymbol{\mu}) \leqslant & -\gamma \underbrace{\sum_{k=0}^{n-1} \sum_{i=1}^{N} \underbrace{m_i \langle \boldsymbol{\mu}^{(i)}(k) - \boldsymbol{\mu}^{(i)}, \ell^{(i)}(\boldsymbol{\mu}(k)) \rangle}_{= \langle \boldsymbol{\mu}^{(i)}(k) - \boldsymbol{\mu}^{(i)}, \nabla_{\boldsymbol{\mu}^{(i)}(k)} \mathrm{V}(\boldsymbol{\mu}(k)) \rangle}}_{\langle \boldsymbol{\mu}(k) - \boldsymbol{\mu}, \mathbf{v}(\boldsymbol{\mu}(k)) \rangle} - \gamma \sum_{k=0}^{n-1} \langle \boldsymbol{\mu}(k) - \boldsymbol{\mu}, \widetilde{\mathcal{M}}^T \boldsymbol{\Lambda}(k) \rangle \\ & + \gamma^2 C_1^2 \sum_{k=0}^{n-1} \|\boldsymbol{\Lambda}(k)\|_2^2 + \gamma S_n + 2\gamma^2 R_n + 2C_2^2 \gamma^2 n \end{aligned} \tag{6.13}$$

where:

$$S_n := -\sum_{k=0}^{n-1} \langle \boldsymbol{X}(k) - x_*, \boldsymbol{M}(k+1) \rangle, \;\; R_n := m_* N \sum_{k=1}^{n} \|\boldsymbol{M}(k)\|_\infty^2,$$

and where $\mathbf{v}(\boldsymbol{\mu}) := \nabla \mathrm{V}(\boldsymbol{\mu})$, where V denotes the Rosenthal potential:

$$\mathrm{V} : \varDelta \to \mathbb{R}, \quad \boldsymbol{\mu} \mapsto \sum_{r \in \mathcal{R}} \int_0^{\boldsymbol{\phi}_r(\boldsymbol{\mu})} \ell_r(u) \mathrm{d}u, \tag{6.14}$$

We now estimate the evolution of the price vector by providing a bound for $\mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) := (\|\boldsymbol{\Lambda}(n) - \boldsymbol{\lambda}\|_2^2 - \|\boldsymbol{\Lambda}(0) - \boldsymbol{\lambda}\|_2^2)/2$, where $\boldsymbol{\lambda} \geqslant 0$. By similar computations as before, and by the elementary bound $2\langle \boldsymbol{\lambda} - \boldsymbol{\Lambda}(k), \boldsymbol{\Lambda}(k) \rangle \leqslant \|\boldsymbol{\lambda}\|_2^2 - \|\boldsymbol{\Lambda}(k)\|_2^2$, we obtain:

$$\begin{aligned} \mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) \leqslant & \, \beta \sum_{k=0}^{n-1} \langle \boldsymbol{\Lambda}(k) - \boldsymbol{\lambda}, \boldsymbol{\phi}(\boldsymbol{\mu}(k)) - L \rangle \\ & + \frac{\alpha}{2} \sum_{k=0}^{n-1} (\|\boldsymbol{\lambda}\|_2^2 - \|\boldsymbol{\Lambda}(k)\|_2^2) + \sum_{k=0}^{n-1} (\beta^2 C_3^2 + \alpha^2 \|\boldsymbol{\Lambda}(k)\|_2^2). \end{aligned} \tag{6.15}$$

Combining the bounds (6.13) and (6.15), , it holds:

$$\begin{aligned}
&\mathcal{V}_n^{(1)}(\boldsymbol{\mu}) + \mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) \\
&\leqslant -\gamma \sum_{k=0}^{n-1} \langle \boldsymbol{z}(k) - \boldsymbol{z}, \tilde{\mathbf{v}}(\boldsymbol{z}(k)) \rangle \\
&+ (\beta - \gamma) \sum_{k=0}^{n-1} \langle \boldsymbol{\Lambda}(k) - \boldsymbol{\lambda}, \widetilde{\mathcal{M}}\boldsymbol{\mu}(k) - L \rangle \\
&+ (\gamma^2 C_1^2 - \frac{\alpha}{2} + \alpha^2) \sum_{k=0}^{n-1} \|\boldsymbol{\Lambda}(k)\|_2^2 \\
&+ \left(2C_2^2\gamma^2 + C_3^2\beta^2 + \frac{\alpha\|\boldsymbol{\lambda}\|_2^2}{2}\right) n + \gamma S_n + 2\gamma^2 R_n,
\end{aligned}$$

where:

$$\begin{aligned}
&\boldsymbol{z}(k) := (\boldsymbol{\mu}(k), \boldsymbol{\Lambda}(k)), \quad z = (\boldsymbol{\mu}, \boldsymbol{\lambda}), \\
&\tilde{\mathbf{v}}(\boldsymbol{z}(k)) = [\nabla \mathrm{V}(\boldsymbol{\mu}(k)) + \widetilde{\mathcal{M}}^T \boldsymbol{\Lambda}(k), L - \widetilde{\mathcal{M}}\boldsymbol{\mu}(k)].
\end{aligned}$$

setting $\beta = \gamma$ and $\alpha = \delta\gamma^2$ with $\delta \in (0, 1/\gamma^2)$ fulfilling (6.6), we have:

$$\begin{aligned}
&\mathcal{V}_n^{(1)}(\boldsymbol{\mu}) + \mathcal{V}_n^{(2)}(\boldsymbol{\lambda}) \\
&\leqslant -\gamma \sum_{k=0}^{n-1} \underbrace{\langle \boldsymbol{z}(k) - \boldsymbol{z}, \tilde{\mathbf{v}}(\boldsymbol{z}(k)) \rangle}_{=:\Upsilon_k(z, \boldsymbol{z}(k))} \\
&+ \left((2C_2^2 + C_3^2)\gamma^2 + \frac{\alpha\|\boldsymbol{\lambda}\|_2^2}{2}\right) n + \gamma S_n + 2\gamma^2 R_n.
\end{aligned} \tag{6.16}$$

Notice that v is monotone since V is convex. Thus by Proposition 4.15, we have that $\tilde{\mathbf{v}}$ is also monotone implying:

$$\Upsilon_k(z, \boldsymbol{z}(k)) \geqslant \langle \boldsymbol{z}(k) - \boldsymbol{z}, \tilde{\mathbf{v}}(z) \rangle.$$

Moreover, by the slater's condition and KKT argumentations, we can find a Lagrangian dual optimizer $\boldsymbol{\lambda}_* \in \mathbb{R}^{\mathcal{R}}_{\geqslant 0}$ corresponding to the minimizer $\boldsymbol{\mu}_*$ of V over $\mathcal{Q} := \{\boldsymbol{\mu} \in \varDelta : \ \boldsymbol{\Gamma}(\boldsymbol{\mu}) \leqslant 0\}$. It follows that $(\boldsymbol{\mu}_*, \boldsymbol{\lambda}_*) \in \mathrm{SOL}(\mathcal{X} \times \mathbb{R}^{\mathcal{R}}, \tilde{\mathbf{v}})$, and consequently:

$$\Upsilon_k(z_*, \boldsymbol{z}(k)) \geqslant \langle \boldsymbol{z}(k) - \boldsymbol{z}_*, \tilde{\mathbf{v}}(z_*) \rangle \geqslant 0, \ z_* = (\boldsymbol{\mu}_*, \boldsymbol{\lambda}_*). \tag{6.17}$$

Setting this observation into (6.16), we obtain:

$$\begin{aligned}
&\mathcal{V}_n^{(1)}(\boldsymbol{\mu}_*) + \mathcal{V}_n^{(2)}(\boldsymbol{\lambda}_*) \\
&\leqslant \left((2C_2^2 + C_3^2)\gamma^2 + \frac{\alpha\|\boldsymbol{\lambda}_*\|_2^2}{2}\right) n + \gamma S_n + 2\gamma^2 R_n.
\end{aligned} \tag{6.18}$$

Now, since $Y_0 = 0$, we have:

$$\mathcal{V}_n^{(1)}(\boldsymbol{\mu}_*) \geqslant -\sum_{i=1}^{N} m_i \left(\max_{\Delta_i} \psi_i - \min_{\Delta_i} \psi_i \right) \geqslant -m_* \sum_{i=1}^{N} \ln(|\mathfrak{P}_i|),$$

and thus $\mathcal{V}_n^{(1)}(\boldsymbol{\mu}_*) \geqslant -\Delta\psi^2/2$. Combining this observation with (6.18) and using $\boldsymbol{\Lambda}_0 = 0$, we obtain that:

$$\begin{aligned}
&\frac{\|\boldsymbol{\Lambda}(n) - \boldsymbol{\lambda}_*\|_2^2}{2} \\
&\leqslant \frac{\Delta\psi^2}{2} + \underbrace{\frac{\|\boldsymbol{\Lambda}(0) - \boldsymbol{\lambda}_*\|_2^2}{2}}_{\frac{\|\boldsymbol{\lambda}_*\|_2^2}{2}} + \left((2C_2^2 + C_3^2)\gamma^2 + \frac{\alpha\|\boldsymbol{\lambda}_*\|_2^2}{2} \right) n \\
&+ \gamma S_n + 2\gamma^2 R_n \\
&= \frac{\Delta\psi^2}{2} + (1 + \alpha n)\frac{\|\boldsymbol{\lambda}_*\|_2^2}{2} + (2C_2^2 + C_3^2)\gamma^2 n \\
&+ \gamma S_n + 2\gamma^2 R_n
\end{aligned} \tag{6.19}$$

Since S_n is a martingale with $\mathbb{E}[S_1] = 0$, we have by taking the expectation (and noticing $\mathbb{E}[S_n] = 0$), the desired result. ■

6.6.2. Proof of consequences of the main result

Proof (Proof of Corollary 6.3): Jensen's and triangle inequality asserts that:

$$\sqrt{\mathbb{E}\left[\|\boldsymbol{\Lambda}(n) - \boldsymbol{\lambda}_*\|_2^2\right]} \geqslant \mathbb{E}\left[\|\boldsymbol{\Lambda}(n) - \boldsymbol{\lambda}_*\|_2\right] \geqslant \mathbb{E}\left[\|\boldsymbol{\Lambda}(n)\|_2\right] - \|\boldsymbol{\lambda}_*\|_2$$

Applying this to (6.7) and by the persistence of the noise, we obtain (6.8).

For any $k \in [n]$, we have by Corollary 6.3:

$$\begin{aligned}
\mathbb{E}\left[\|\boldsymbol{\Lambda}(k)\|_2\right] \leqslant &\Delta\psi + (1 + \sqrt{(1 + \delta\gamma^2 n)})\|\boldsymbol{\lambda}_*\|_2 \\
&+ (\tilde{C}_1 + \sigma)\gamma\sqrt{n}.
\end{aligned}$$

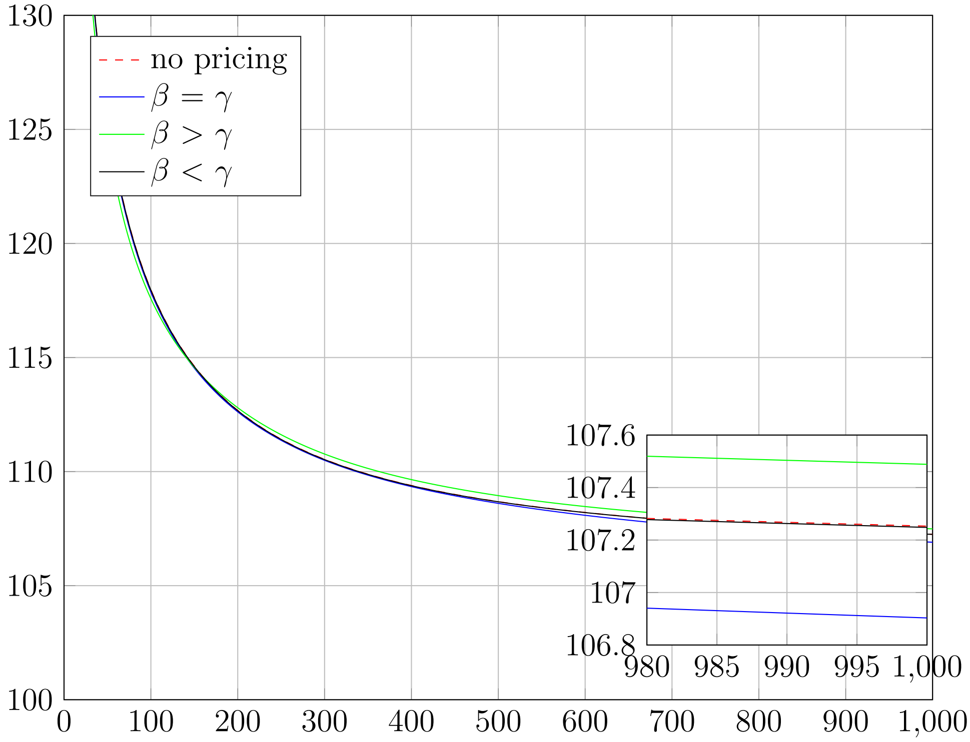
Now, setting our choices of parameters into (6.8), it yields:

$$\begin{aligned}
\mathbb{E}\left[\|\boldsymbol{\Lambda}(k)\|_2\right] \leqslant &\Delta\psi + (1 + \sqrt{(1 + \delta c^2)})\|\boldsymbol{\lambda}_*\|_2 \\
&+ (\tilde{C}_1 + \sigma)c = A.
\end{aligned}$$

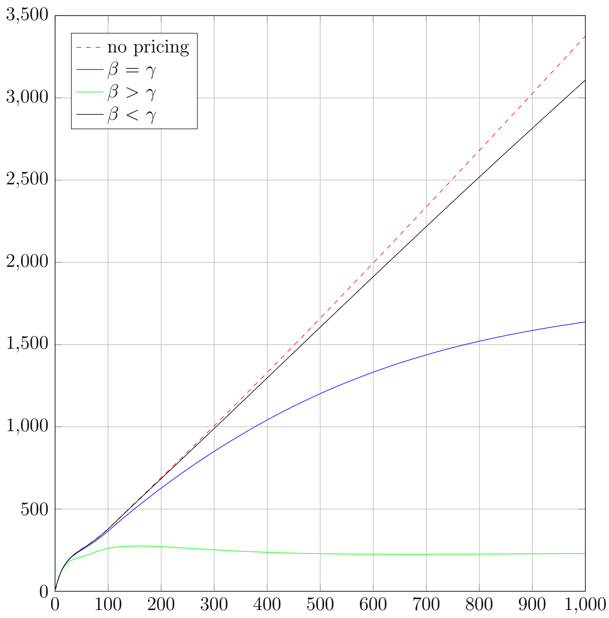
Consequently:

$$\begin{aligned}\frac{\alpha}{\beta}\mathbb{E}\left[\sum_{k=0}^{n-1}\|\boldsymbol{\Lambda}(k)\|_2\right] &= \frac{\delta c}{\sqrt{n}}\sum_{k=1}^{n-1}\mathbb{E}\left[\|\boldsymbol{\Lambda}(k)\|_2\right] \\ &\leqslant \frac{\delta c(n-1)}{\sqrt{n}}A \leqslant \delta c A\sqrt{n}.\end{aligned} \tag{6.20}$$

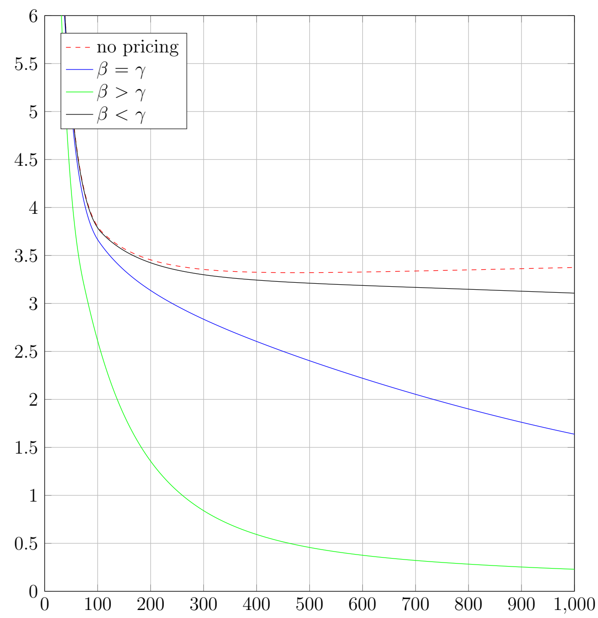
Moreover, we have $\mathbb{E}\left[\|\boldsymbol{\Lambda}(n)\|_2\right]/\beta \leqslant A\sqrt{n}/c$. Setting this observation and (6.20) into (6.4), we have the remaining statement. ∎



(a) AD averaged over time

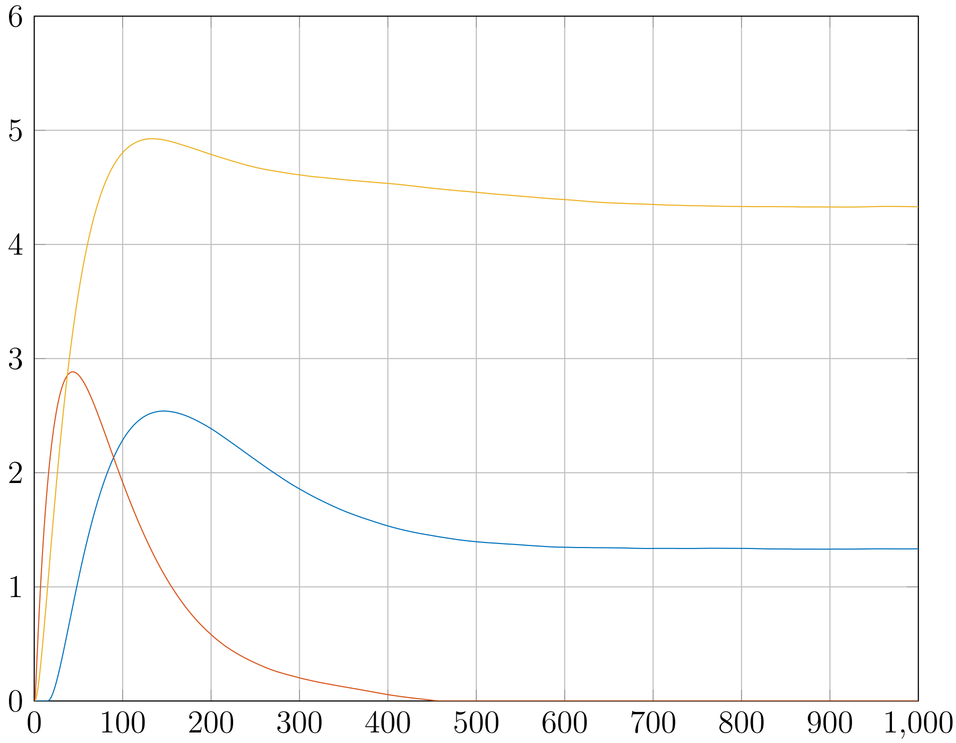


(b) ACV

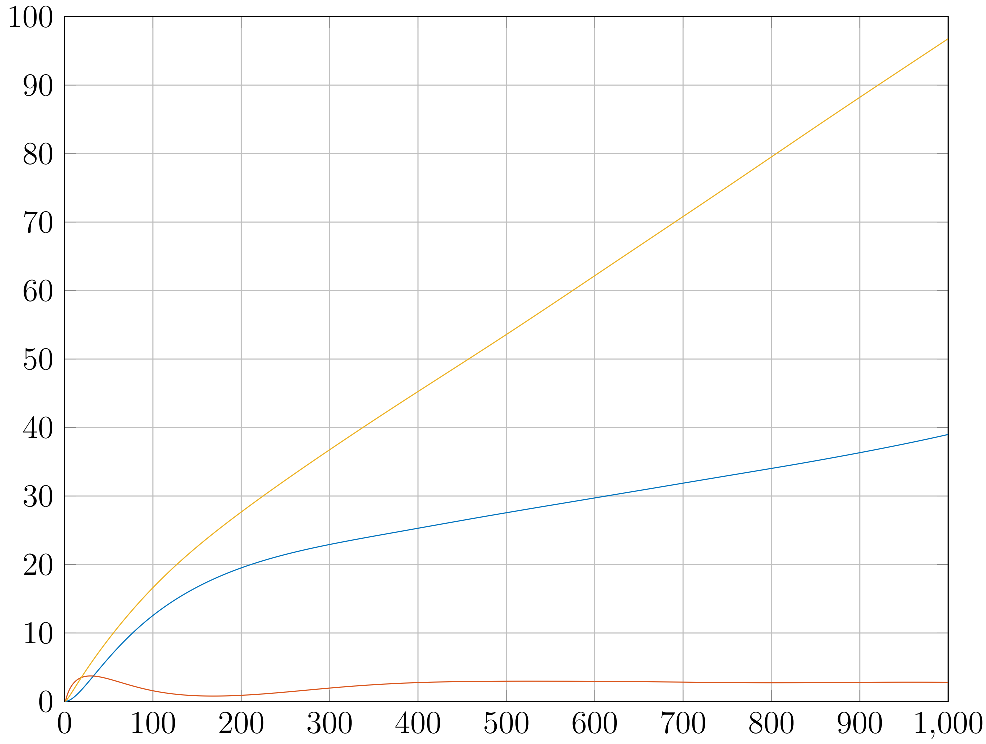


(c) ACV averaged over time

Figure 6.3.: Performance for $L_r = 14$

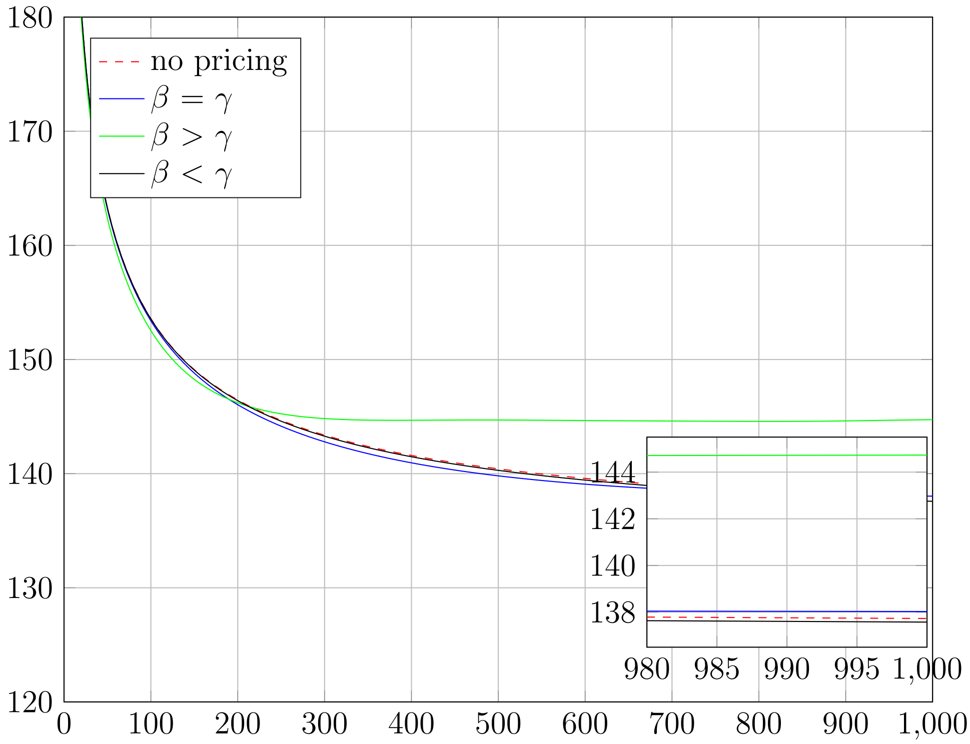


(a) $L_r = 14$

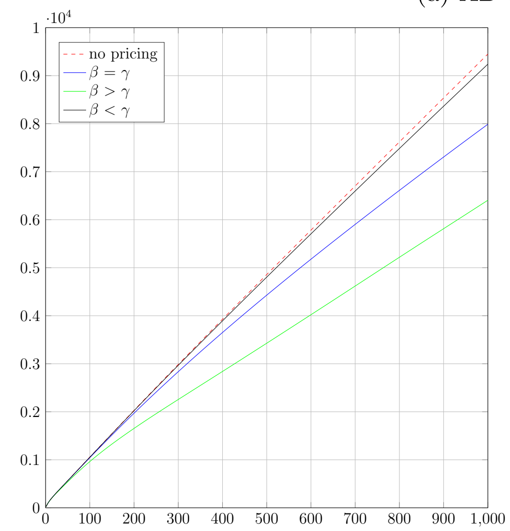


(b) $L_r = 11$

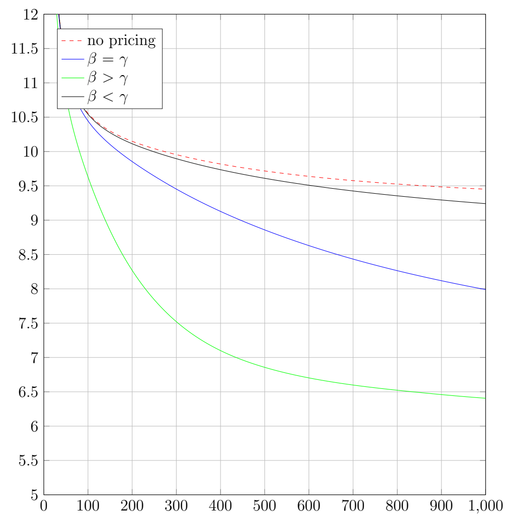
Figure 6.4.: Pricing over time



(a) AD averaged over time



(b) ACV



(c) ACV averaged over time

Figure 6.5.: Performance for $L_r = 11$

Part II.

Online Decision-Making with Emphasize on the Noise-Robustness and Sustainable Behaviour

7. Robust Online Learning for Resource Allocation - Beyond Euclidean Projection and Dynamic Fit

Abstract: Online-learning literature has focused on designing algorithms that ensure sub-linear growth of the cumulative long-term constraint violations. The drawback of this guarantee is that strictly feasible actions may cancel out constraint violations on other time slots. For this reason, we introduce a new performance measure called h-CFit, whose particular instance is the cumulative positive part of the constraint violations. We propose a class of non-causal algorithms for online-decision making, which guarantees, in slowly changing environments, sub-linear growth of this quantity despite noisy first-order feedback. Furthermore, we demonstrate by numerical experiments the performance gain of our method relative to the state of art.

7.1. Introduction

Classical OL deal with problems with time-invariants constraints that has to be strictly satisfied. Therefore, a projection operator is typically applied to the update. However, in practical applications (see e.g. [30, 35–37]) one usually encounter additional time-variant constraints. Moreover, the need for decentralization of the learner action in applications can not be satisfied by simply using a centralized projection operator. For those reasons, several works [30, 35, 111, 131, 132] propose projected primal-dual methods which ensure sub-linear growth of the regret, i.e., the cumulative distance of the generated action to the optimal ones, and the long-term fulfillment of the constraints in the sense that the sum of the constraint violations grow sub-linearly. A problem relates to this long-term guarantee is that it holds, despite substantial instantaneous constraint violations, as long as the methods generate strictly feasible actions canceling the latter.

Our Contribution In this work, we introduce a new long-term constraint preservation performance measure called h-CFit. The feature of this performance measure is that h-CFit avoids cancellation effects between the summands, which might occur in the simple cumulative constraint violation measure. We design a non-causal saddle-point method based on mirror descent aiming to ensure dynamic regret optimization and sublinear growth of h-CFit. In particular, we can guarantee dynamic regret bound of order $\mathcal{O}(\mathbb{V}_T T^{1/2})$, where $\mathbb{V}_T$ measures the variation of the optimizers of the underlying time-varying problem, and h-CFit-bound of order $\mathcal{O}(T^{3/4})$. We show by numerical experiments the performance gain of our method relative to state of the art and the advantage of using a mirror map other than Euclidean projection.

Table 7.1.: An Overview of Related Works on Online Convex Optimization

References	Long-Term Constraint Type	Feedback Noise	Regret Bound
[25]	No	No	$\mathcal{O}(T^{1/2})$
[111, 132, 133]	$\sum_{t=1}^{T} \mathbf{g}(x_t)$	No	$\mathcal{O}(T^{1/2})$
[134]	$\sum_{t=1}^{T} \mathbf{g}_t(x_t)$	No	$\mathcal{O}(T^{1/2})$
[131]	$\sum_{t=1}^{T}[\mathbf{g}(x_t)]_+^2$	No	$\mathcal{O}(T^{1/2})$
This paper	$\sum_{t=1}^{T} \mathrm{h}(\mathbf{g}_t(\boldsymbol{X}_t))$,	Martingale	$\mathcal{O}((1+\sigma^2+\mathbb{V}_T)T^{1/2})$

References	Regret Benchmark Type	Constraint Violation Bound
[25]	Static and Dynamic	-
[111, 132, 133]	Static	$\mathcal{O}(T^{3/4})$ resp. $\mathcal{O}(T^{1/2})$
[134]	Static	$\mathcal{O}(T^{1/2})$
[131]	Static	$\mathcal{O}(T^{1/2})$
This paper	Dynamic	$\mathcal{O}(T^{3/4})$

References	Comments
[25]	Mirror Map
[111, 132, 133]	Mirror Map
[132, 134]	Requires Slater condition, Causal dual update, Euclidean projection
[131]	
This paper	Mirror Map

Relation to Prior Works In the absence of long-term constraints, [25] showed that the standard method of online mirror descent achieves $\mathcal{O}(\sqrt{T})$ regret bound (see also [38]), which is known to be optimal [43]. However, their notion of regret, i.e., static regret, corresponds to the difference of the losses between the online solution and the overall best static solution in hindsight, which is weaker than ours.

The first work tackling the online problem with long-term constraints is [111]. This work considers time-invariant constraint function and proposed an algorithm having $\mathcal{O}(\sqrt{T})$ regret bound, and $\mathcal{O}(T^{3/4})$ cumulative constraint violation bound. As investigated by [133], one can efficiently trade-off between those bound by allowing the step-size to be variable. By utilizing Slater's condition, and allowing the dual update to depend causally on the primal update, [132] provides an improved $\mathcal{O}(\sqrt{T})$ bound for the cumulative constraint violation. [131] was able to provide $\mathcal{O}(\sqrt{T})$ bound for tighter long-term constraint preservation measure $\sum_{t=1}^{T}[\mathbf{g}(\boldsymbol{X}_t)]_+^2$. However, this remarkable guarantee is achieved by allowing the dual update to utilize causal information about the primal variable.

To the best of our knowledge, the first work considering the online problem with time-varying constraints is [134]. Based on [132], they provide a (causal) primal-dual algorithm ensuring that the static regret is of order $\mathcal{O}(\sqrt{T})$, and the cumulative constraint violation of order $\mathcal{O}(\sqrt{T})$.

Until now, we only discuss works delivering static regret guarantees. The work [30] proposed a projected gradient descent based algorithm for the online problem with time-varying constraints aiming to optimize the regret against the dynamic comparator. Their result relies on the assumption that two consecutive constraint functions are bounded by the slack achieved by a fixed primal action uniformly over all constraint functions. Surely both, the existence of the slack and the action, and the boundedness of the difference consecutive constraint functions are difficult to guarantee. Despite this fact, their dynamic regret bound is worse than ours ($\mathcal{O}(\mathbb{V}_T T^{1/2})$) since it is lower bounded by $\mathcal{O}(\mathbb{V}_T T^{1/3})$. The cumulative constraint violations guarantee given [30] is of order $\mathcal{O}(T^{2/3})$, which is better than ours ($\mathcal{O}(T^{3/4})$). However, our performance measure to this respect, i.e., h-CFit, is stronger than that given in [30]. A clear plus-point of [30] is the application of the proposed online algorithm to proactive network allocation. [35] proposed a novel adaptive algorithm for the online problem with time-varying constraints with interesting applications to the problem of computational offloading in IoT. The corresponding method possesses higher computational complexity than ours since it requires the covariance of the gradient, root, and inverse operation of a matrix. Despite of this fact, their dynamic regret guarantee ($\mathcal{O}(T^{7/8}\mathbb{V}_T)$ and long-term constraint preservation guarantee ($\mathcal{O}(\max\left\{T^{15/16}, T^{7/8}\sqrt{\mathbb{V}_T}\right\})$) is worse than ours.

Basic Notions and Notations For a real vector a, $[a]_+$ denotes the vector whose entries are the non-negative part of the entries of a. The canonical projection onto a closed convex subset A of an Euclidean space $\mathbb{R}^D$ is denoted by $\mathbf{\Pi}_A$, i.e.:

$$\mathbf{\Pi}_A(\boldsymbol{y}) := \underset{\boldsymbol{x}\in A}{\arg\min}\,\|\boldsymbol{y}-\boldsymbol{x}\|_2.$$

For a subspace A of an Euclidean space $\mathbb{R}^D$, we denote the diameter of A by:

$$D_A := \sup_{\boldsymbol{x},\boldsymbol{y}\in A}\|\boldsymbol{x}-\boldsymbol{y}\|_2$$

Let $(\mathcal{X}, \|\cdot\|)$ be a normed space and $A, B \subseteq \mathcal{X}$. We denote $A - B := \{\boldsymbol{x} - y : \ x \in A,\ y \in B\}$, and $\|A\| := \sup_{\boldsymbol{x}\in A}\|\boldsymbol{x}\|$. In this work we assume that a probability space $(\Omega, \Sigma, \mathbb{P})$ and a filtration $\mathbb{F} := (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ therein are given.

7.2. Problem Formulation

We begin by stating the online learning problem in the classical setting:

Online Learning with Classical Aggregate constraint goal

At each time t, a learner decides for an option for action $\boldsymbol{X}_t$ from an apriori known compact convex set $\mathcal{X} \subset \mathbb{R}^D$, which we refer to as a *feasible set*. Subsequently, nature chooses the *loss function* $f_t^{(i)}$ and charges the learner with loss $\mathrm{f}_t(\boldsymbol{X}_t)$. As already noticed in prior works, it is advantageous from a practical point of view to take into account a time-varying *penalty function* $\mathbf{g}_t = (\mathrm{g}_t^r)_{r\in[R]}$ chosen by nature and revealed to the learner at a time t. This function leads to a time-varying *constraint* $\mathbf{g}_t(\boldsymbol{X}_t) \leqslant 0$. In an online learning setting, one often assumes additionally that the learner can extract information about f_t and $\mathbf{g}_t$ via access to the first-order oracle in order to choose the action for the time step $t+1$. Although this assumption is sometimes not realistic, investigation respective to this case is usually a stepping stone for designing methods in the bandit case, i.e., in the case where the learner has only access to the immediate objective- and constraint value (see e.g. Chapter 4 in [38] and [111]).

Given a time horizon $T \in \mathbb{N}$. The goal of the learner is to find a sequence $(\boldsymbol{X}_t)_{t=1}^T$ in the feasible set $\mathcal{X}$ that minimizes the loss $\mathrm{f}_t(\boldsymbol{X}_t)$ and simultaneously fulfills the constraint $\mathbf{g}_t(\boldsymbol{X}_t) \leqslant 0$. Since the learner cannot look into the future and therefore has to decide on her next action utilizing the current information about the loss and penalty function, the problem stated before is intractable. For this reason, one may consider a more realistic goal of finding a sequence that minimizes the time-average loss:

$$\sum_{t=1}^{T} \mathrm{f}_t(\boldsymbol{X}_t)/T,$$

and that ensures the fulfillment of the constraint on average over time

$$\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{X}_t)/T \leqslant 0. \tag{7.1}$$

Beyond Aggregate Constraint

One crucial issue about the latter goal concerning the constraint fulfillment (7.1) is that it does not consider the possibility that the summands can cancel each other out: As long as $\mathbf{g}_t(\boldsymbol{X}_t)$ is negative and small enough for specific time slots $t \in [T]$, large $\mathbf{g}_t(\boldsymbol{X}_t)$ for another time slots $t \in [T]$ is admissible for the goal $\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{X}_t)/T \leqslant 0$. In order to resolute this issue, we propose a new online learning goal, that is:

$$\min_{(\boldsymbol{X}_t)_{t=1}^T \subset \mathcal{X}} \sum_{t=1}^{T} \mathrm{f}_t(\boldsymbol{X}_t) \quad \text{s.t.} \quad \sum_{t=1}^{T} \mathrm{h}(\mathbf{g}_t(\boldsymbol{X}_t)) \leqslant 0 \tag{7.2}$$

for a monotonically increasing function $h : \mathbb{R} \to \mathbb{R}$. In case that h is also non-negative, this function ensures that cancellation between summands cannot occur since they are all non-negative and that the ordering between values of $\mathbf{g}_t$ remains preserved.

An example of h is:

$$\mathrm{h}(\cdot) = [\cdot]_+$$

. This choice leads to the constraint:

$$\sum_{t=1}^{T} [\mathbf{g}_t(x_t)]_+ / T \leqslant 0$$

, that is stronger than (7.1). Another example of h is:

$$\mathrm{h}(\cdot) = [\cdot]_+^p, \quad p > 1,$$

leading to the constraint:

$$\sum_{t=1}^{T} [\mathbf{g}_t(x_t)]_+^p / T \leqslant 0$$

. With increasing p, $[\cdot]_+^p$ penalizes large values of $\mathbf{g}_t$ with the cost of loosening the sensitivity of the sum for small, non-negative values of $\mathbf{g}_t$.

Noisy First-order Feedback

As discussed in Subsection 7.2, the online learning setting assumes that first-order information about the current loss function is available at each time slot. However, perfect first-order feedback is, in general, hard to obtain. Thus, we include in our model the possibility that the learner has only access to the noisy first-order oracle. Expressly, we assume that at each time t and for a given action $\boldsymbol{X}_t \in \mathcal{X}$, the learner can query an estimate of $\hat{\mathbf{v}}_t$ of the (sub-)gradient $\nabla \mathrm{f}_t(\boldsymbol{X}_t)$ satisfying:

$$\mathbb{E}[\|\hat{\mathbf{v}}_t\|_*] < \infty \quad \text{and} \quad \mathbb{E}[\hat{\mathbf{v}}_t | \mathcal{F}_t] = \nabla f_t(\boldsymbol{X}_t),$$

where $\mathcal{F}_t$ is an element of a filtration $\mathbb{F} := (\mathcal{F}_t)_{t \in \mathbb{N}_0}$ on a probability space $(\Omega, \Sigma, \mathbb{P})$. The canonical and commonly-used filtration in the literature is the filtration of the history of the considered iterates. Equivalently, we can model the stochastic (sub-)gradient by

$$\hat{\mathbf{v}}_t = \nabla \mathrm{f}(\boldsymbol{X}_t) + \boldsymbol{M}_{t+1}, \tag{7.3}$$

where $(\boldsymbol{M}_t)_{t\in\mathbb{N}}$ is a $\mathbb{R}^D$-valued $\mathbb{F}$-*martingale difference sequence*, i.e. it is $\mathbb{F}$-*adapted*, i.e.:

- $\boldsymbol{M}_t$ is $\mathcal{F}_t$-measureable for all $t \in \mathbb{N}$
- its members are *conditionally mean zero*, in the sense that:

$$\mathbb{E}[\boldsymbol{M}_t|\mathcal{F}_{t-1}] = 0, \quad \text{for all } t \in \mathbb{N}.$$

7.2.1. Applications

In order to show the practical relevance of the the aspects discussed above (especially: noisy feedback and other notion of aggregate constraint), we give in the following some specific resource allocation examples.

Example 13 (Economic Dispatch): Consider a system with D producers (e.g. electric generator or data processing center) of a certain commodity (e.g. electrical power or data processing unit). At each time slot $t \in \mathbb{N}$, the goal of economic dispatch is to decide for each $i \in [D]$ the output $\boldsymbol{X}_t^{(i)}$ of producer i causing costs $\mathrm{c}_t^{(i)}(\boldsymbol{X}_t^{(i)})$ such that

$$\text{the total producing cost:} \quad \sum_{i=1}^{D} \mathrm{c}_t^{(i)}(\boldsymbol{X}_t^{(i)}),$$

remains low, and the extrinsic given

$$\text{demand:} \quad d_t$$

is balanced. A possible loss function to this regard is:

$$\mathrm{f}_t : \mathbb{R}^D \to \mathbb{R},\ \boldsymbol{x} \mapsto \sum_{i=1}^{D} \mathrm{c}_t^{(i)}(\boldsymbol{x}^{(i)}) + \xi \left(\sum_{i=1}^{D} \boldsymbol{x}^{(i)} - d_t \right)^2,$$

where $\xi > 0$

In solving the economic dispatch problem, one has to consider several constraints. For instance, the output of each producer $i \in [D]$ can not exceed the value $x_{\max}^{(i)} \in \mathbb{R}_{\geqslant 0}$ specified e.g. technical restrictions of the producer. Since the violation of this constraint might not be tolerable, Prior works settle the feasible set in the online learning problem formulation as the box-type set

$$\mathcal{X} = \left\{ \boldsymbol{x} \in \mathbb{R}^D :\ 0 \leqslant \boldsymbol{x}^{(i)} \leqslant \boldsymbol{x}_{\max}^{(i)},\ \forall i \in [D] \right\}.$$

This kind of feasible set is popular in applications (see e.g. [26,30]). Instead of considering the constraint specified by technical restrictions of the producers, one may instead consider the constraint specified by total output production resulting in the feasible set:

$$\mathcal{X} = \left\{ \boldsymbol{x} \in \mathbb{R}^D_{\geqslant 0} : \sum_{i=1}^{D} \boldsymbol{x}^{(i)} \leqslant B \right\}.$$

In the application of economic dispatch for electrical power, above feasible set corresponds to power transmission restriction specified by the wireline capacity.

Another constraint which one may consider is that the total negative externality (e.g. pollution) $\sum_{i=1}^{D} E_t^{i\to j}(x_t^{(i)})$ of a certain kind (e.g. substance) $j \in [N]$ should not exceed a particular value $E^j_{\max}$ (e.g. specified by government regulator). In the previous sum, $E_t^{i\to j}$ denotes a function specifying the negative externality of a certain kind $j \in [N]$ at time slot t given a specific output of the producer $i \in [D]$. This gives rise to the penalty function:

$$\mathbf{g}_t : \mathbb{R}^D \to \mathbb{R}^N,\ \boldsymbol{x} \mapsto \left(\sum_{i=1}^{D} E_t^{i\to j}(\boldsymbol{x}^{(i)}) - E^j_{\max} \right)_{j=1}^{N} \tag{7.4}$$

In the strict sense, it is absurd to think that inter-time compensation of negative externalities occurs. For instance, pollution causes damages irrespective of whether in the earlier time emission constraint is strictly preserved. Rather, the system manager should ensure that $\mathbf{g}_t(\boldsymbol{X}_t)$ remains at each time step small. For this reason, the aim of preserving the relaxed constraint given in (7.2) seems to be more plausible than the aim of preserving $\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{X}_t)/T \leqslant 0$.

In order to see where disturbance of the gradient feedback might occur, let us assume that the cost $\mathrm{c}_t^{(i)}$ is quadratic, i.e.:

$$\mathrm{c}_t^{(i)}(x) = \mathbf{a}_t^{(i)} x^2 + \boldsymbol{b}_t^{(i)} x,$$

where $\mathbf{a}_t^{(i)}$ and $\boldsymbol{b}_t^{(i)}$ are non-negative constants depending on the specific sort of producer. For instance, if the considered commodity is the electrical power and the considered producer is a steam turbine unit, the constants $\mathbf{a}_t^{(i)}$ and $\boldsymbol{b}_t^{(i)}$ depend on the current fuel price, changing over time, and on the maintenance price, including labor price [135]. The first-order information of the cost function can be given explicitly as $\nabla \mathrm{f}_t(x) = (\mathbf{a}_t^{(i)} x + \boldsymbol{b}_t^{(i)})_i$. In reality, one usually has only a disturbed observation of the prices $\mathbf{a}_t^{(i)}$ and $\boldsymbol{b}_t^{(i)}$. For instance, considered the previous instance, the disturbance is due to the uncertainty of the estimate of the current fuel cost. We model this fact by defining the noisy feedback as follows:

$$\hat{\mathbf{v}}_t^{(i)} = (\mathbf{a}_t^{(i)} + \tilde{\boldsymbol{M}}_{t+1}^{(i),1})\boldsymbol{X}_t + \boldsymbol{b}_t^{(i)} + \tilde{\boldsymbol{M}}_{t+1}^{(i),2},$$

where $(\tilde{\boldsymbol{M}}_t^{(i),1})_t$ and $(\tilde{\boldsymbol{M}}_t^{(i),2})_t$ are martingale w.r.t. a filtration containing the history of $(\boldsymbol{X}_t)_t$. It holds:

$$\mathbb{E}_t[\tilde{\boldsymbol{M}}_{t+1}\boldsymbol{X}_t] = \mathbb{E}_t[\tilde{\boldsymbol{M}}_{t+1}]\boldsymbol{X}_t = 0,$$

and thus by defining:

$$\boldsymbol{M}_{t+1}^{(i)} = \tilde{\boldsymbol{M}}_{t+1}^{(i),1}\boldsymbol{X}_t + \tilde{\boldsymbol{M}}_{t+1}^{(i),2}$$

it follows that $\hat{\mathbf{v}}_t = \nabla \mathrm{f}_t(\boldsymbol{X}_t) + \boldsymbol{M}_{t+1}$, where $(\boldsymbol{M}_t)$ is a martingale. This formulation of $\hat{\mathbf{v}}_t$ coincides with the model described in Subsection 7.2.

Example 14 (Trajectory Tracking): Consider a dynamical system:

$$\boldsymbol{X}_{t+1} = \mathbf{A}\boldsymbol{X}_t + \mathbf{B}\boldsymbol{u}_t,$$

where $\boldsymbol{X}_t$ is the location of a robot and $\boldsymbol{u}_t$ is the control action. Let $\boldsymbol{Y}_t$ be the location of the target at time slot t. The objective of trajectory tracking at time slot t is to choose a control action $\boldsymbol{u}_t$ s.t. the tracking error:

$$\mathrm{f}_t(\boldsymbol{X}_t) = \frac{\|\boldsymbol{X}_t - \mathbf{Y}_t\|_2^2}{2}$$

and

$$\text{the smoothness measure:} \quad (\beta/2)\|\boldsymbol{X}_t - \boldsymbol{X}_{t-1}\|_2^2, \quad \text{where } \beta > 0,$$

is minimized. Possible constraint which one may consider is the energy constraint $\|u_t\|_2^2 \leqslant u_{2,\max}$ and extremum control value constraints $u_{\min} \leqslant u_t^{(i)} \leqslant u_{\max}$. Considering a time horizon T we may solve for a given initial states x_0, the following online problem:

$$\begin{aligned} &\min_{(\boldsymbol{u}_t)_{t=1}^T} \sum_{t=1}^{T} \|\mathbf{A}\boldsymbol{x}_t + \mathbf{B}\boldsymbol{u}_t - \boldsymbol{y}_{t+1}\|_2^2 + \frac{\beta}{2}\|(\mathbf{A}-\mathbf{I})x_t + \mathbf{B}\boldsymbol{u}_t\|_2^2 \\ &\text{s.t.} \ \sum_{i=1}^{D} (\boldsymbol{u}_t^{(i)})^2 \leqslant u_{2,\max} \\ &\qquad \boldsymbol{u}_{\min}^{(i)} \leqslant \boldsymbol{u}_t^{(i)} \leqslant \boldsymbol{u}_{\max}^{(i)} \quad i \in [D]. \end{aligned} \tag{7.5}$$

Defining the loss function as:

$$\mathrm{f}_t(\boldsymbol{u}) = \|\mathbf{A}\boldsymbol{x}_t + \mathbf{B}u - \boldsymbol{y}_{t+1}\|_2^2 + \frac{\beta}{2}\|(\mathbf{A}-\mathbf{I})x_t + \mathbf{B}\boldsymbol{u}\|_2^2,$$

we obtain that the loss feedback is given by:

$$\nabla \mathrm{f}_t(\boldsymbol{u}) = 2\mathbf{B}^{\mathrm{T}}(\mathbf{A}\boldsymbol{x}_t - \boldsymbol{y}_{t+1}) + (2+\beta)\mathbf{B}^{\mathrm{T}}\mathbf{B}\boldsymbol{u} + \beta\left[\mathbf{B}^{\mathrm{T}}(\mathbf{A}-\mathbf{I})\boldsymbol{x}_t\right].$$

A possible source of disturbance in the gradient feedback is the location $\boldsymbol{y}_{t+1}$ of the target at time $t+1$. One may also consider the sparsity constraint $\|u_t\|_1 \leqslant u_{1,\max}$ instead of/in addition to the energy constraint.

Example 15 (Energy Harvesting Communications): Suppose there are N communication systems (CS) where the transmitters uses the power from B power sources (PS) in order to transmit over a time-varying channel. For a bundle $(\boldsymbol{X}_t^{i\to j})_i$ of power allocated by the power sources to CS $j \in [N]$ at time t, we assume the (ideal) transmit rate of this CS obey the relationship:

$$r_t^j = \mu_t^j \log_2 \left(1 + \left[\sum_{i=1}^{B} a_t^{i\to j} \boldsymbol{X}_t^{i\to j}\right] g_t^j\right),$$

where μ_t^j denotes the bandwidth of the channel used by CS j, $g_t^{(j)}$ denotes the instantaneous gain-to-noise ratio of the channel used by CS j, and $a_t^{i\to j} \in [0,1]$ denotes the factor of loss incurred by sending the amount of power from PS i to CS j. Let p_t^i be the amount of power harvested by the power source $i \in [B]$. Suppose that the PS $i \in [B]$ possesses a battery in order to store the excess energy for further use at the subsequent time $t+1$. The state of PS i's battery at time $t+1$ is given by:

$$Q_{t+1}^i = Q_t^i + p_t^i - \sum_{j=1}^{N} \boldsymbol{X}_t^{i\to j},$$

where negative Q_t^i means that PS i needs at time t to import costly energy from external sources which. So one constraint for the online strategy is $Q_t^i \geqslant 0 \quad \forall t \in [T]$ which is basically:

$$-\sum_{\tau=1}^{t} \left(p_t^i - \sum_{j=1}^{N} \boldsymbol{X}_t^{i\to j}\right) \leqslant 0, \quad \forall t \in [T-1].$$

Another constraint which we may consider is the overcharging constraint i.e. the state of the battery at time t has to be below a certain power treshhold:

$$-\left(p_t^i - \sum_{j=1}^{N} \boldsymbol{X}_t^{i\to j}\right) \leqslant Q_{\max}, \quad \forall t \in [T-1].$$

The reason that one may see overcharging constraint as a long-term constraint is that overcharging is admissible.

$$\max_{(x_t)_{t=1}^T} \sum_{t=1}^{T}\sum_{i=1}^{N} B_t^{(i)} \log_2 \left(1 + \left[\sum_{j=1}^{B} a_t^{j\to i} x_t^{(j)}\right] g_t^{(i)}\right)$$

$$\sum_{j=1}^{N} x_t^{i\to j} \leqslant p_t^{(i)} \quad \text{(Service Constraint)}$$

$$p_t^{(i)} - \sum_{j} a_t^{i\to j} x_t^{(i\to j)} \geqslant Q_{\max}^{(j)} \quad \text{(Overcharging constraint)}$$

Example 16 (Online Network Resource Allocation): We consider the resource allocation problem over a cloud network (see [30, 136]) represented by a directed graph $\mathcal{G} = (\mathcal{I}, \mathcal{E})$, where $\mathcal{I} = \mathcal{J} \cup \mathcal{K}$ denotes the set of nodes containing data centers (DC) $\mathcal{K}$ and mapping nodes (MN) $\mathcal{J}$, and where $\mathcal{E}$ denotes the edge set containing all the links between MN and DC.

At each time slot t every MN $j \in \mathcal{J}$ receives a data processing request d_t^j. At the same time slot, each MN j may forwards the amount $\boldsymbol{X}_t^{j\to k}$ to DC $k \in \mathcal{K}$. In the same time slot, each DC can schedules workload of amount $\boldsymbol{X}_t^k$ for processing. We may model the cost of applying a data allocation decision $\boldsymbol{X}_t = ((\boldsymbol{X}_t^{i\to j})_{j\in\mathcal{K},i\in\mathcal{I}},\dots,(\boldsymbol{X}_t^{I\to j})_{j\in\mathcal{K}},(\boldsymbol{X}_t^j)_{j\in\mathcal{K}})$ as follows:

$$\begin{aligned}\mathrm{f}_t(\boldsymbol{X}_t) &= \sum_{k\in\mathcal{K}} \underbrace{\mathrm{f}_t^k(\boldsymbol{X}_t^k)}_{\text{Power cost}} + \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} \underbrace{\mathrm{f}_t^{j\to k}(\boldsymbol{X}_t^{j\to k})}_{\text{Bandwidth Cost}} \\ &+ \xi_1 \sum_{j\in\mathcal{J}} \underbrace{\left(d_t^i - \sum_{j\in\mathcal{K}} \boldsymbol{X}_t^{i\to j}\right)^2}_{\text{Demand Service}} + \xi_2 \sum_{k\in\mathcal{K}} \underbrace{\left(\boldsymbol{X}_t^k - \sum_{i\in\mathcal{I}} \boldsymbol{X}_t^{i\to k}\right)^2}_{\text{Processing Service}}.\end{aligned}$$

Above cost structure contains the computational resources needed by a DC $k \in \mathcal{K}$ in order to process the amount $\boldsymbol{X}_t^k$ of tasks and the communication resources needed by MN $j \in \mathcal{J}$ in order to forward the amount $\boldsymbol{X}_t^{j\to k}$ of tasks to the data center k. The difference between above cost structure to the one given in the previous works [30, 136] is that the former considers additionally the cost incurred by not forwarding the instantaneous demand to the processing unit and the cost incurred by ignoring the current task queue.

which differs from the previous work. One may consider uploading limit:

$$\sum_{k\in\mathcal{K}} X^{j\to k} \leqslant \boldsymbol{x}_{\max}^k, \quad \forall j,$$

where $\boldsymbol{x}_{\max}^k$ is a specific constant

7.2.2. Performance measure and Our Goal

In this work, we use the following performance measure, called dynamic regret (see [25]), which is defined for a sequence of decisions $\boldsymbol{X}_1,\dots,\boldsymbol{X}_t$ of the online learner as follows:

$$\mathrm{Reg}_t^{\mathrm{d}} := \sum_{\tau=1}^{t} (\mathrm{f}_\tau(\boldsymbol{X}_\tau) - \mathrm{f}_\tau(x_\tau^*)) \quad \text{(Dynamic Regret)},$$

where our benchmark is the sequence of the best dynamic solution $\boldsymbol{x}_t^* := (x_\tau^*)_{\tau\in[t]}$ with:

$$x_\tau^* \in \arg\min_{\boldsymbol{x}\in\mathcal{X}} \mathrm{f}_\tau(x) \text{ s.t. } \mathbf{g}_\tau(x) \leqslant 0, \quad \forall \tau \in [t].$$

Throughout this work, we assume that the following regularity condition on the cost function holds:

Assumption 7.1: • f_τ *is convex and subdifferentiable of* $\mathcal{X}$.

- *For each* $\tau \in [t]$ *and for all* $x \in \mathcal{X}$, *we have a fix choice of subgradient* $\nabla \mathrm{f}_\tau(x)$ *such* *that* $\sup_{\boldsymbol{x}\in\mathcal{X}} \|\nabla \mathrm{f}_\tau(x)\|_* < \infty$.

In general $\mathrm{Reg}_t^{\mathrm{d}}$ can be negative. This case occurs if for some $\tau \in [t]$, $\boldsymbol{X}_\tau$ is not feasible w.r.t. to the constraint $\mathbf{g}_\tau(x) \leqslant 0$. However, we have the following lower bound by the mean value Theorem:

$$\mathrm{Reg}_t^{\mathrm{d}} \geqslant -\sum_{\tau=1}^{t} \sup_{\mathcal{X}_\tau} \|\nabla \mathrm{f}_\tau\|_* \|\boldsymbol{x}_\tau - \boldsymbol{x}_\tau^*\| \geqslant -D_{\mathcal{X}} \sum_{\tau=1}^{t} L_\tau, \tag{7.6}$$

where $L_\tau > 0$ is a constant fulfilling:

$$\|\nabla \mathrm{f}_\tau\|_* \leqslant L_\tau.$$

Related to the dynamic regret, is the following performance measure called dynamic gap defined for $\boldsymbol{x}_t := (x_\tau)_{\tau\in[t]} \subset \mathcal{X}$ as follows:

$$\mathrm{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t) := \sum_{\tau=1}^{t} \langle \boldsymbol{X}_\tau - \boldsymbol{x}_\tau, \nabla \mathrm{f}_\tau(\boldsymbol{X}_\tau)\rangle,$$

which we often use in our analysis. The reason is that besides:

$$\mathrm{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t^*) \geqslant \mathrm{Reg}_t^{\mathrm{d}}, \tag{7.7}$$

which follows from the convexity of f_τ, for all $\tau \in [T]$, the gradients which constitute the building blocks of PDOGA appears in its formulation.

Performance measure for the feasibility of the learner decision respective to the constraint $\mathbf{g}_\tau(x) \leqslant 0$, $\tau \in [t]$, which we use in this work is the following:

$$\text{h-CFit}_t^r := \sum_{\tau=1}^{t} \mathrm{h}(\mathbf{g}_r(\boldsymbol{X}_t)).$$

We assume that the following regularity condition on h-CFit:

Assumption 7.2:

- *For all $r \in [N]$, $\mathbf{g}_t^{(r)}$ is convex and (sub-)differentiable on $\mathcal{X}_t$.*
- *h is monotonically increasing and sub-differentiable on $\mathbb{R}$.*

7.3. Algorithm Design

In this section, we provide a novel algorithm which we call generalized online mirror saddle-point (GOMSP) whose aim is to generate online decisions minimizing the performance measures introduced in Subsection 7.2.2. For convenience, we provide a summary of our finding in Algorithm 7.

Algorithm 7 Generalized Online Mirror Saddle-Point (GOMSP) Method

Require: Time horizon $T \in \mathbb{N}$, learning rate $\gamma > 0$, price sensitivity $\beta > 0$, regularization constant $\alpha > 0$.

Require: Initial score $\boldsymbol{Y}_1 \in \mathbb{R}^D$, - primal iterate $\boldsymbol{X}_1 = \boldsymbol{\Phi}(\boldsymbol{Y}_1)$, - dual variable $\boldsymbol{\Lambda}_1 \in \mathbb{R}^R_{\geqslant 0}$

for $t = 0, 1, 2, \ldots, T$ **do**

Observe the noisy first-order feedback

$$\hat{\mathbf{v}}_t := \nabla \mathrm{f}_t(\boldsymbol{X}_t) + \boldsymbol{M}_{t+1}$$

for $r = 1, \ldots, R$ **do**

Query the first-order h-load feedback $\nabla(\mathrm{h} \circ \mathbf{g}_t^r)(\boldsymbol{X}_t)$

end for

Update the score vector as in (7.8)

Update the dual variable:

$$\boldsymbol{\Lambda}_{t+1} = \boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}} \left[(1 - \alpha\gamma)\boldsymbol{\Lambda}_t + \gamma \mathrm{h}(\mathbf{g}_t(\boldsymbol{X}_t))\right]$$

Update primal variable as in (7.9):

end for

7.3.1. Primal Variable Update - Mirror Descent

The basis of the primal update of GOMSP is the score vector which is generated from the actual noisy first-order objective - and constraint feedback by the following rule:

$$\boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t - \gamma \left(\hat{\mathbf{v}}_t + \sum_{r=1}^{N} \left[\nabla(\mathrm{h} \circ \mathbf{g}_t^r)(\boldsymbol{X}_t)\right] \boldsymbol{\Lambda}_t^{(r)} \right). \tag{7.8}$$

The variable $\boldsymbol{\Lambda}_t^{(r)}$ is a Lagrange variable that corresponds to the r-th constraint, whose update rule will be specified later.

To realize the primal update $\boldsymbol{X}_{t+1}$ from the score vector $\boldsymbol{Y}_{t+1}$ at the time slot $t+1$, we use the mirror map $\boldsymbol{\Phi}$ (see Definition 2.1). Clearly, the mirror map is a generalization of the usual Euclidean projection. An interesting example of mirror maps is the so-called logit choice $\boldsymbol{\Phi}(y) = \exp(y)/\sum_{l=1}^{D} \exp(y_l)$ which is generated by the 1-strongly convex regularizer $\psi(x) = \sum_{k=1}^{D} x_k \log x_k$ on the probability simplex $\Delta \subset (\mathbb{R}^D, \|\cdot\|_1)$. Other instance of mirror map worth to mentions is $\boldsymbol{\Phi}(Y) = \exp(Y)/(1 + \|\exp(Y)\|_1)$ which is defined on the set $\mathcal{X}$ of positive semidefinite matrices X having the nuclear norm $\|X\|_1 := \mathrm{tr}(|X|) \leqslant 1$. The von-Neumann entropy $\psi(\boldsymbol{X}) = \mathrm{tr}(X \log X) + (1 - \mathrm{tr}X)\log(1 - \mathrm{tr}X)$ is a (1/2)-strongly convex regularizer on $\mathcal{Z}$ [45] (for derivation see e.g. [46]).

Having introduced the notion of the mirror map, we can define the primal update rule given a score vector $\boldsymbol{Y}_{t+1}$ and a regularizer ψ as follows:

$$\boldsymbol{X}_{t+1} = \boldsymbol{\Phi}(\boldsymbol{Y}_{t+1}). \tag{7.9}$$

In case that the chosen regularizer is the Euclidean norm, one can write:

$$\boldsymbol{X}_{t+1} = \boldsymbol{\Pi}_{\mathcal{X}}\left[\boldsymbol{X}_t - \gamma\left(\hat{\mathbf{v}}_t + \sum_{r=1}^{N}\left[\nabla(\mathrm{h}\circ\mathbf{g}_t^r)(\boldsymbol{X}_t)\right]\boldsymbol{\Lambda}_t^{(r)}\right)\right],$$

which is the update rule for the projected noisy gradient descent related to the online Lagrangian:

$$\mathcal{L}_t(x,\lambda) = \mathrm{f}_t(x) + \lambda^{\mathrm{T}}\mathrm{h}(\mathbf{g}_t(x))), \quad x \in \mathcal{X},\ \lambda \in \mathbb{R}^N_{\geqslant 0}. \tag{7.10}$$

This Lagrangian corresponds to the optimization problem:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \mathrm{f}_t(x) \quad \text{s.t.} \quad \mathrm{h}(\mathbf{g}_t(x)) \leqslant 0.$$

This observation explains our motivation for defining the primal update as in (7.8) and (7.9) in case the underlying projection operator is Euclidean.

The reason to use a "projection" mapping, which is in our case the mirror map, more general than Euclidean projection is that it yields a versatile method for the online decision-making process. The mirror map allows us to adapt the first-order penalized iterative method to the geometry of the underlying feasible set of the decision problem and to leverage from the weaker dimension dependency of the algorithm performance. This effect has been recognized earlier in connection with the simple gradient descent method [47, 48]: Using the logit choice instead of Euclidean projection for realizing iterative simple first-order descent method for convex optimization problem on simplex yields a convergence guarantee which depends logarithmically on the - instead of the square root of the underlying dimension. Moreover, using a mirror map other than the Euclidean projection might yield a better dimension dependency of the noise term in the resulted bound since the noise influence is no longer measured by the Euclidean norm.

Another factor that is variable in the update rule of GOMSP is the function h. In this regard, we provide for the convenience of the reader a particular form of (7.8) in the following:

Example 17: The function h which we mainly have in mind is $\mathrm{h}(\cdot) = [\cdot]_+^p$. By choosing the subgradient as follows:

$$\nabla(\mathrm{h}\circ\mathbf{g}_t^r)(x) = \begin{cases} p\ (\mathbf{g}_t^r(x))^{p-1}\nabla\mathbf{g}_t^{(r)}(x) & \text{if } \mathbf{g}_t^{(r)}(x) \geqslant 0 \\ 0 & \text{else,} \end{cases}$$

we may write:

$$\boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t - \gamma\left(\hat{\mathbf{v}}_t + p\sum_{r\in\mathcal{A}_t}(\mathbf{g}_t^r(\boldsymbol{X}_t))^{p-1}\nabla\mathbf{g}_t^r(\boldsymbol{X}_t)\boldsymbol{\Lambda}_t^{(r)}\right),$$

where $\mathcal{A}_t = \{r \in [R] :\ \mathbf{g}_t^r(x) > 0\}$ denotes the set of active constraints at time t.

7.3.2. Dual variable update

The primary role of the dual variable $\boldsymbol{\Lambda}_t$ is to provide the primal variable information about the actual amount of the constraint violation. One might draw the analogy between this variable and the prices in markets whose role is to signal the participants to what extent the corresponding resources are scarce. In particular, $\boldsymbol{\Lambda}_t$ has to reflect the actual constraint violation state. Besides, another crucial requirement for the dual variable is that it does not grow unboundedly. Otherwise, the constraint term in the primal update overthrow the cost part and consequently the primal update concentrates on reducing the amount of violation rather than minimizing the regret.

In hindsight of those aspects, we give the update rule the dual variable of GOMSP as follows:

$$\boldsymbol{\Lambda}_{t+1} = \mathbf{\Pi}_{\mathbb{R}^R_{\geqslant 0}} \left[(1-\alpha\gamma)\boldsymbol{\Lambda}_t + \gamma \mathrm{h}(\mathbf{g}_t(\boldsymbol{X}_t))\right]. \tag{7.11}$$

In case that $\alpha = 0$, (7.11) turn to the simple dual gradient ascent corresponds to the Lagrangian (7.10). The idea behind adding the regularization term $\alpha\gamma\boldsymbol{\Lambda}_t$ is to reduce the growth of the dual variable by decaying the influence of previous constraint states: To see this, notice that if $h \geqslant 0$, we can omit the projection operator in the expression (7.11). Consequently:

$$\boldsymbol{\Lambda}_{t+1} = (1-\alpha\gamma)^t \boldsymbol{\Lambda}_1 + \gamma \sum_{\tau=1}^{t} (1-\alpha\gamma)^{\tau-t} \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau)).$$

If $\boldsymbol{\Lambda}_1 = 0$, we have:

$$\boldsymbol{\Lambda}_{t+1} = \gamma \sum_{\tau=1}^{t} (1-\alpha\gamma)^{t-\tau} \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau)).$$

Thus the influence of the τ-th constraint function term to the dual variable at time $t+1$ decays with $\exp(-\tau a)$ where $a = -\ln(1-\alpha\gamma)$. This can be advantageous in the online environment since $\mathbf{g}_\tau$ for different time-slots (with large distance) not necessarily correlate.

Another reason for defining (7.11) is that the resulted dual dynamic gives rise about the cumulative constraint state in the following sense:

Lemma 7.1 (From Dual Dynamic to Constraint Violation): *Suppose that $\boldsymbol{\Lambda}_1 = 0$. It holds:*

$$\text{h-CFit}_t^r \leqslant \frac{\|\boldsymbol{\Lambda}_{t+1}\|_2}{\gamma} + \alpha \sum_{\tau=1}^{t} \|\boldsymbol{\Lambda}_\tau\|_2$$

Proof: By (7.11), we have $\boldsymbol{\Lambda}^r_{\tau+1} \geqslant \boldsymbol{\Lambda}^r_\tau + \gamma \mathrm{h}(g^r_\tau(\boldsymbol{X}_\tau)) - \alpha\gamma\boldsymbol{\Lambda}^r_\tau$. So summing, telescoping, and the assumption $\boldsymbol{\Lambda}_1 = 0$ give:

$$\gamma \sum_{\tau=1}^{t} \mathrm{h}(g^r_\tau(\boldsymbol{X}_\tau)) \leqslant \boldsymbol{\Lambda}^r_{t+1} + \alpha\gamma \sum_{\tau=1}^{t} \boldsymbol{\Lambda}^r_\tau \tag{7.12}$$

By the fact that $\boldsymbol{\Lambda}^r_t \geqslant 0$, it holds $\boldsymbol{\Lambda}^r_t \leqslant \|\boldsymbol{\Lambda}_t\|_2$. Finally, the latter and (7.12) give the desired inequality. ∎

7.4. Performance Analysis

To analyze the performance of the algorithm, we leverage from Lyapunov-type argumentation. In doing that, we use as energy functions both, the distance between the iterate $\boldsymbol{Y}_t$ of the algorithm and the current constraint minimizer of the cost function and the norm of the dual variable:

$$\mathcal{E}_t(x) = \underbrace{\mathrm{F}(x, \boldsymbol{Y}_t)}_{=:\mathcal{E}_t^1(x)} + \underbrace{\frac{\|\boldsymbol{\Lambda}_t\|_2}{2}}_{=:\mathcal{E}_t^2},$$

where F denotes the Fenchel coupling (see Definition 2.2). This sort of Lyapunov function is standard (see e.g. [25]) besides the fact that we use the Fenchel coupling as the primal iterate distance function.

To analyze the performance of the proposed algorithm, we give in the following an upper bound for the primal dynamic and dual dynamic.

7.4.1. Lyapunov Analysis

Primal Dynamic For convenience, we rewrite (7.8) as:

$$\boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t - \gamma\left(\hat{\mathbf{v}}_t - [\nabla(\mathrm{h}\circ\mathbf{g}_t)(\boldsymbol{X}_t)]^{\mathrm{T}}\boldsymbol{\Lambda}_t\right),$$

where:

$$[\nabla(\mathrm{h}\circ\mathbf{g}_t)(\boldsymbol{X}_t)]^{\mathrm{T}} = \left[\nabla(\mathrm{h}\circ g_t^1)(\boldsymbol{X}_t), \ldots, \nabla(\mathrm{h}\circ\mathbf{g}_t^R)(\boldsymbol{X}_t)\right]$$

The following result gives the upper bound of the one-step difference $\Delta\mathcal{E}_t^1(x) := \mathcal{E}_{t+1}^1(x) - \mathcal{E}_t^1(x)$:

Lemma 7.2: *For any $x \in \mathcal{X}$:*

$$\begin{aligned}\Delta\mathcal{E}_t^1(x) \leqslant & -\gamma\langle\boldsymbol{X}_t - \boldsymbol{x}, \nabla\mathrm{f}_t(\boldsymbol{X}_t)\rangle \\ & -\gamma\langle\boldsymbol{X}_t - \boldsymbol{x}, [\nabla(\mathrm{h}\circ\mathbf{g}_t)(\boldsymbol{X}_t)]^{\mathrm{T}}\boldsymbol{\Lambda}_t\rangle \\ & +\gamma\tilde{\boldsymbol{M}}_{t+1} + \frac{\gamma^2 C_{1,\psi}^2}{K}\|\boldsymbol{\Lambda}_t\|_2^2 + \frac{2\gamma^2}{K}(C_{2,\psi}^2 + \|\boldsymbol{M}_{t+1}\|_*^2),\end{aligned}$$

where $C_{1,\psi}, C_{2,\psi}$, are the smallest constants $C_1, C_2 > 0$ satisfying for all $\lambda \in \mathbb{R}_{\geqslant 0}^R$ and $x \in \mathcal{X}$:

$$\|[\nabla(\mathrm{h}\circ\mathbf{g}_t)(x)]^{\mathrm{T}}\lambda\|_* \leqslant C_1\|\lambda\|_2 \quad \|\nabla\mathrm{f}_t(x)\|_* \leqslant C_2$$

Dual Dynamic The expression given in Lemma 7.2 possesses a dependency on the dual variable $\boldsymbol{\Lambda}_t$. So to continue, it stands clear to analyze the dynamic of this variable. Toward this direction, we have the following result on the drift of the Lagrangian:

Lemma 7.3: *For $\tilde{\boldsymbol{x}} \in \mathcal{Q}_\tau$:*

$$\Delta\mathcal{E}_\tau^{(2)} \leqslant \gamma\langle[\nabla(\mathrm{h}\circ\mathbf{g}_\tau)(\boldsymbol{X}_\tau)]^{\mathrm{T}}\boldsymbol{\Lambda}_\tau, \boldsymbol{X}_\tau - \tilde{\boldsymbol{x}}\rangle - (\alpha\gamma - \alpha^2\gamma^2)\|\boldsymbol{\Lambda}_\tau\|_2^2 + \gamma^2 C_3^2,$$

where $C_3 > 0$ is a constant satisfying:

$$\|\mathrm{h}(\mathbf{g}(x))\|_2 \leqslant C_3, \quad \forall x \in \mathcal{X}. \tag{7.13}$$

For ease of the readibility, we provide the proof of this Lemma in Appendix 7.7.1.

Primal-Dual Dynamic By combining previous auxiliary statements on the dynamic of the primal - and dual variable, we obtain the following result:

Theorem 7.4: *Suppose that:*

$$\alpha - \gamma(\alpha^2 - \frac{C_1^2}{K}) \geqslant 0 \tag{7.14}$$

For $\boldsymbol{x}_t := (x_\tau)_{\tau\in[t]} \subset \mathcal{X}$ with $x_\tau \in \mathcal{Q}_\tau$ for all $\tau \in [t]$:

$$\mathrm{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t) + \frac{\|\boldsymbol{\Lambda}_{t+1}\|_2^2}{2\gamma} \leqslant -\frac{\mathcal{V}_t(\boldsymbol{x}_t)}{\gamma} + \frac{\|\boldsymbol{\Lambda}_1\|_2^2}{2\gamma} + t\gamma C_\psi^2 + S_t(\boldsymbol{x}_t) + \frac{2\gamma}{K}R_t,$$

where:

$$S_t(\mathbb{u}_t) = \sum_{\tau=1}^{t}\tilde{\boldsymbol{M}}_{\tau+1}(u_\tau), \quad R_t = \sum_{\tau=1}^{t}\|\boldsymbol{M}_{\tau+1}\|_*^2,$$
$$\mathcal{V}_t^1(\boldsymbol{x}_t) := \sum_{\tau=1}^{t}\Delta\mathcal{E}_\tau^1(x_\tau), \quad C_\psi^2 := \frac{2C_{2,\psi}^2}{K} + C_{3,\psi}^2,$$

Proof: From Lemma 7.2 and Lemma 7.3, we obtain for any $\tilde{\boldsymbol{x}} \in \mathcal{Q}_t$:

$$\begin{aligned}
&\Delta\mathcal{E}_t^1(\tilde{\boldsymbol{x}}) + \Delta\mathcal{E}_t^2 \\
&\leqslant -\gamma\langle\boldsymbol{X}_t - \tilde{\boldsymbol{x}}, \nabla\mathrm{f}_t(\boldsymbol{X}_t)\rangle + \gamma\tilde{\boldsymbol{M}}_{t+1} + \frac{2\gamma^2}{K}(C_2^2 + \|\boldsymbol{M}_{t+1}\|_*^2) \\
&\underbrace{-\gamma(\alpha - \gamma\alpha^2 - \frac{\gamma C_1^2}{K})\|\boldsymbol{\Lambda}_t\|_2^2}_{=:e_1} + \gamma^2 C_3^2.
\end{aligned}$$

The condition (7.14) help us to get rid of the expression $e1$, which involves the dual variable. By summing the resulted inequality and since $2\sum_{\tau=1}^{t}\Delta\mathcal{E}_t^2 = \|\boldsymbol{\Lambda}_{t+1}\|_2^2 - \|\boldsymbol{\Lambda}_1\|_2^2$, we obtain the desired statement. ∎

By the relation (7.7), $\mathrm{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t^*)$ gives rise to the dynamic regret. Moreover, Lemma 7.1 asserts that the Lagrangian variable contains the information about the cumulation of the constraint violation. Thus we come closer to achieving the objective of providing performance guarantee for the proposed algorithm. As usual, the terms S_t and R_t due to objective feedback noise can be handled by taking the expectation. So, the only term at which a closer look should be taken is $\mathcal{V}_t^1(\boldsymbol{x}_t^*)$.

Lower bound for Primal Energy Function In case that that the environment is not adversary, i.e., f_τ remains for all $\tau \in [t]$ the same, it holds by telescoping:

$$\mathcal{V}_t^1(\boldsymbol{x}_t^*) = \mathrm{F}(x^*, \boldsymbol{Y}_{t+1}) - \mathrm{F}(x^*, \boldsymbol{Y}_1) \geqslant -\mathrm{F}(x^*, \boldsymbol{Y}_1),$$

where x^* denotes the constrained minimizer of f_τ. What we may do in the adversary case is to interpolate $\mathcal{V}_t^1(\boldsymbol{x}_t^*)$ by the cumulative difference of the benchmark sequence $\boldsymbol{x}_t$. In order to execute this procedure, we assume the following:

Assumption 7.3: *The regularizer is nowhere steep in the sense that ψ is differentiable on $\mathcal{X}$.*

Before we proceed, we first discuss this assumption in the following:

Remark 20: Suppose that $\mathcal{X} = \left\{\boldsymbol{x} \in \mathbb{R}^D_{\geqslant 0} : \sum_{i=1}^D x_i \leqslant B\right\}$ for a fixed constant $B > 0$. The Euclidean norm seen as a regularizer on $\mathcal{X}$ is clearly nowhere steep. In contrast to the Euclidean norm, the entropy function $\psi(x) = \sum_{i=1}^D x_i \ln(x_i)$ as a regularizer is not nowhere steep since the gradient of ψ grows unboundedly as the argument goes to the element of $\mathcal{X}$ which possesses zero coordinates. However, we may instead use the smoothed entropy $\psi_\epsilon(x) = \psi(x+\epsilon)$ where $\epsilon > 0$ is a chosen constant. As we will discuss later This procedure does not have any significant impact on the dynamic of our algorithm.

We first show that is the regularizer is nowhere steep then the Fenchel coupling is Lipschitz in the first argument:

Lemma 7.5: *Suppose that ψ is nowhere steep. Then for all $x_1, x_2 \in \mathcal{X}$ and $y \in \mathbb{R}^D$:*

$$|F_\psi(x_1, y) - F_\psi(x_2, y)| \leqslant 2L_\psi \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|,$$

where $L_\psi > 0$ is given by:

$$L_\psi := \sup_{\boldsymbol{x} \in \mathcal{X}} \|\nabla\psi(x)\|_*. \tag{7.15}$$

Proof: By definition of F and the triangle inequality, we have:

$$|\mathrm{F}(x_1, y) - \mathrm{F}(x_2, y)| \leqslant |\psi(x_1) - \psi(x_2)| + |\langle y, \boldsymbol{x}_1 - \boldsymbol{x}_2\rangle|$$

Mean value Theorem and the nowhere-steepness of ψ asserts:

$$|\psi(x_1) - \psi(x_2)| \leqslant L_\psi \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|.$$

Now, since ψ is nowhere steep, it follows from Proposition 2.3 that $\boldsymbol{\Phi}$ is surjective. So we can find a $x \in \mathcal{X}$ s.t. $x = \mathbf{\Phi}(y)$ and thus (again by Proposition 2.3) $y = \nabla\psi(x)$. Consequently we have by Hölder inequality:

$$|\langle y, \boldsymbol{x}_1 - \boldsymbol{x}_2 \rangle| \leqslant \|\nabla\psi(x)\|_* \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leqslant L_\psi \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|. \qquad \blacksquare$$

We are now ready to give a lower bound for $\mathcal{V}_t^1$:

Lemma 7.6: *Suppose that* ψ *is nowhere steep. It holds:*

$$\mathcal{V}_t^{(1)}(\boldsymbol{x}_t) \geqslant -\mathrm{F}(x_1, \boldsymbol{Y}_1) - L_\psi \mathbb{V}^\psi(\boldsymbol{x}_{t+1}),$$

where L_ψ *is given in* (7.15)*:*

$$\mathbb{V}^\psi(\boldsymbol{x}_{t+1}) := \sum_{\tau=1}^{t} \|\boldsymbol{x}_{\tau+1}^* - \boldsymbol{x}_\tau^*\|$$

Proof: Applying Lemma 7.5, we obtain:

$$\begin{aligned}
\Delta\mathcal{E}_\tau^{(1)}(x_\tau) &= \mathrm{F}(x_\tau, \boldsymbol{Y}_{\tau+1}) - \mathrm{F}(x_\tau, \boldsymbol{Y}_\tau) \\
&= \mathrm{F}(x_{\tau+1}, \boldsymbol{Y}_{\tau+1}) - \mathrm{F}(x_\tau, \boldsymbol{Y}_\tau) \\
&\quad + \mathrm{F}(x_\tau, \boldsymbol{Y}_{\tau+1}) - \mathrm{F}(x_{\tau+1}, \boldsymbol{Y}_{\tau+1}) \\
&\geqslant \mathrm{F}(x_{\tau+1}, \boldsymbol{Y}_{\tau+1}) - \mathrm{F}(x_\tau, \boldsymbol{Y}_\tau) - L_\psi \|\boldsymbol{x}_{\tau+1} - \boldsymbol{x}_\tau\|
\end{aligned}$$

Thus summing over $\tau \in [t]$:

$$\begin{aligned}
\mathcal{V}_t^{(1)}(\boldsymbol{x}_{t+1}) &\geqslant \mathrm{F}(x_{t+1}, \boldsymbol{Y}_{t+1}) - \mathrm{F}(x_1, \boldsymbol{Y}_1) - L_\psi \mathbb{V}_t(\boldsymbol{x}_{t+1}) \\
&\geqslant -\mathrm{F}(x_1, \boldsymbol{Y}_1) - L_\psi \mathbb{V}_t(\boldsymbol{x}_{t+1})
\end{aligned} \qquad \blacksquare$$

7.4.2. Dynamic Regret bound

Theorem 7.7: *Suppose that* ψ *is nowhere steep and that* (7.14) *is fulfilled. For* $\mathbb{u}_t := (u_\tau)_{\tau\in[t]} \subset \mathcal{X}$ *with* $u_\tau \in \mathcal{Q}_\tau$ *for all* $\tau \in [t]$*:*

$$\mathbb{E}[\mathrm{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t)] \leqslant \frac{\mathrm{F}(x_1, \boldsymbol{Y}_1)}{\gamma} + L_\psi \frac{\mathbb{V}(\boldsymbol{x}_t)}{\gamma} + \frac{\|\boldsymbol{\Lambda}_1\|_2^2}{2\gamma} + t\gamma C_\psi^2 + \frac{2\gamma}{K} \sum_{\tau=1}^{t} \sigma_{t+1}^2,$$

where:

$$C_\psi^2 := \frac{2C_{2,\psi}^2}{K} + C_{3,\psi}^2,$$

and:

$$\mathbb{E}[\|\boldsymbol{M}_\tau\|_*^2] \leqslant \sigma_\tau^2.$$

Proof: First notice that $\|\boldsymbol{\Lambda}_t\|_2 \geqslant 0$. Combining this with (7.14), it holds:

$$\mathrm{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t) \leqslant \frac{\mathrm{F}(x_1^*, \boldsymbol{Y}_1)}{\gamma} + \frac{L_\psi \mathbb{V}_t}{\gamma} + \frac{\|\boldsymbol{\Lambda}_1\|_2^2}{2\gamma} + \gamma t C_\psi^2 + S_t(\boldsymbol{x}_t) + \frac{2\gamma}{K} R_t. \tag{7.16}$$

Now, one can check that $S_t(\boldsymbol{x}_t)$, $t \in \mathbb{N}$ is a martingale. Consequently:

$$\mathbb{E}[S_t(\boldsymbol{x}_t)] = \mathbb{E}[\langle \boldsymbol{X}_1 - \boldsymbol{x}_1, \boldsymbol{M}_2 \rangle] = \mathbb{E}[\langle \boldsymbol{X}_1 - \boldsymbol{x}_1, \mathbb{E}[\boldsymbol{M}_2 | \mathcal{F}_1] \rangle] = 0. \qquad \blacksquare$$

So taking the expectation over (7.16), we obtain the desired statement.

Corollary 7.8: *Suppose that the requirements of Theorem 7.7 are fulfilled and suppose that in addition $\boldsymbol{Y}_1 = 0$ and $\boldsymbol{\Lambda}_0 = 0$. Moreover suppose that the noise is persistent in the sense that there exists $\sigma > 0$ s.t. $\mathbb{E}[\|\boldsymbol{M}_\tau\|_*^2] \leqslant \sigma^2$ for all τ. With:*

$$\gamma = \Theta(T^{-1/2}),$$

it holds:

$$\mathbb{E}[\mathrm{Reg}_t^{\mathrm{d}}] \leqslant [\mathcal{D}(\mathcal{X}, \psi) + L_\psi \mathbb{V}_t]\, \mathcal{O}(\sqrt{T}) + (C_\psi^2 + \frac{\sigma^2}{K}) t \mathcal{O}(T^{-1/2})$$

where:

$$\mathcal{D}(\mathcal{X}, \psi) = \sup_{\boldsymbol{x} \in \mathcal{X}} \psi(x) - \inf_{\boldsymbol{x} \in \mathcal{X}} \psi(x) \quad C_\psi^2 = C_{2,\psi}^2 + C_{3,\psi}^2$$

Proof: By the relation (7.7), it follows that the upper bound for the gap given By the assumption $\boldsymbol{Y}_1 = 0$, it holds:

$$\mathrm{F}(x_1, \boldsymbol{Y}_1) = \psi(x_1) - \psi^*(0) = \psi(x_1) - \inf_{\boldsymbol{x} \in \mathcal{X}} \psi(x) \leqslant \mathcal{D}(\mathcal{X}, \psi)$$

Previous observations and the assumption $\boldsymbol{\Lambda}_1 = 0$ and the asummption that the noise is persistent yields:

$$\mathbb{E}[\mathrm{Reg}_t^{\mathrm{d}}] \leqslant \frac{\mathcal{D}(\mathcal{X}, \psi) + L_\psi \mathbb{V} t}{\gamma} + \left(C_\psi^2 + \frac{2\sigma^2}{K} \right) t\gamma, \qquad \blacksquare$$

So from above result, we have that $\mathbb{E}[\mathrm{Reg}_T^{\mathrm{d}}]$ is of order $\mathcal{O}((1 + \mathbb{V}_t + \sigma)\sqrt{T})$ in case that the online environment changes slowly in the sense that $\mathbb{V}_T \leqslant \mathcal{O}(T^p)$ where $p < 1/2$, the expected regret is sublinear.

7.4.3. Constraint Violation Analysis

Requirements:

$$\|\nabla \mathrm{f}_t(x)\|_* \leqslant L_f, \quad \forall x \in \mathcal{X}, t \tag{7.17}$$

$$\bigcap_{\tau \in [t]} \mathcal{Q}_\tau \neq \varnothing.$$

(7.17) asserts that for $x \in \mathcal{X}$:

$$-\langle \boldsymbol{X}_t - \boldsymbol{x}, \nabla \mathrm{f}_t(\boldsymbol{X}_t)\rangle \leqslant \|\boldsymbol{X}_t - \boldsymbol{x}\| \|\partial \mathrm{f}_t(\boldsymbol{X}_t)\|_* \leqslant D_{\mathcal{X}} L_f$$

Theorem 7.9: *For any $\tilde{\boldsymbol{x}} \in \bigcap_{\tau \in [t]} \{\mathbf{g}_\tau \leqslant 0\}$, it holds:*

$$\frac{\mathbb{E}[\|\boldsymbol{\Lambda}_{t+1}\|_2^2]}{2} \leqslant \gamma t \mathcal{D}_{\mathcal{X}} L_f + F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_1) + \frac{\|\boldsymbol{\Lambda}_1\|_2^2}{2} + t\gamma^2 C_\psi^2 + \frac{2\gamma^2 \sum_{\tau=1}^t \sigma_{\tau+1}^2}{K}$$

Proof: We have for any $\boldsymbol{x}_t \subset \mathcal{X}$:

$$-\operatorname{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t) = -\sum_{\tau=1}^{t} \langle \boldsymbol{X}_\tau - \boldsymbol{x}_\tau, \nabla \mathrm{f}_\tau(\boldsymbol{X}_\tau)\rangle \leqslant t D_{\mathcal{X}} L_f,$$

and for $\boldsymbol{x}_t \subset \mathcal{X}$ with $x_\tau = \tilde{\boldsymbol{x}} \in \mathcal{X}$ for all τ:

$$\mathcal{V}_t^1(\boldsymbol{x}_t) = F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_{t+1}) - F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_1) \geqslant -F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_1)$$

Combining this with Theorem 7.4, it holds for $\boldsymbol{x}_t \subset \mathcal{X}$ with $x_\tau = \tilde{\boldsymbol{x}} \in \bigcap_\tau \mathcal{Q}_\tau$ for all τ:

$$\begin{aligned}\frac{\mathbb{E}[\|\boldsymbol{\Lambda}_{t+1}\|_2^2]}{2} &\leqslant -\gamma \mathbb{E}[\operatorname{Gap}_t^{\mathrm{d}}(\boldsymbol{x}_t)] + F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_1) + \frac{\|\boldsymbol{\Lambda}_1\|_2^2}{2} + t\gamma^2 C_\psi^2 + \frac{2\gamma^2 \sum_{\tau=1}^t \sigma_{\tau+1}^2}{K}, \\ &\leqslant \gamma t \mathcal{D}_{\mathcal{X}} L_f + F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_1) + \frac{\|\boldsymbol{\Lambda}_1\|_2^2}{2} + t\gamma^2 C_\psi^2 + \frac{2\gamma^2 \sum_{\tau=1}^t \sigma_{\tau+1}^2}{K}\end{aligned}$$

■

Corollary 7.10: *Suppose that $\boldsymbol{Y}_1 = 0$ and $\boldsymbol{\Lambda}_1 = 0$. For $\gamma = \Theta(T^{-1/2})$ and $\alpha = \Theta(T^{-1/2})$ fulfilling* (7.14), *It holds:*

$$\text{h-CFit}_t^r \leqslant \sqrt{\mathcal{D}(\mathcal{X}, \psi)}\mathcal{O}(T^{1/2}) + \sqrt{D_{\mathcal{X}} L_f}\mathcal{O}(T^{3/4}) + (C_\psi^2 + \frac{2\sigma^2}{K})^{1/2}\mathcal{O}(T^{1/2})$$

Proof: By the assumption $\boldsymbol{Y}_1 = 0$, we have $F(\tilde{\boldsymbol{x}}, \boldsymbol{Y}_1) \geqslant \mathcal{D}(\mathcal{X}, \psi)$. So, it holds:

$$\frac{\mathbb{E}[\|\boldsymbol{\Lambda}_{t+1}\|_2^2]}{2} \leqslant \mathcal{D}_{\mathcal{X}} L_f t \mathcal{O}(T^{-1/2}) + \mathcal{D}(\mathcal{X}, \psi) + (C^2 + \frac{2\sigma^2}{K}) t \mathcal{O}(T^{-1})$$

Consequently by Jensen's inequality:

$$\mathbb{E}[\|\boldsymbol{\Lambda}_{t+1}\|_2] \leqslant \sqrt{\mathcal{D}(\mathcal{X}, \psi)} + \sqrt{D_{\mathcal{X}} L_f} t^{1/2} \mathcal{O}(T^{-1/4}) + (C_\psi^2 + \frac{2\sigma^2}{K})^{1/2} t^{1/2} \mathcal{O}(T^{-1/2}).$$

Consequently:

$$\frac{\mathbb{E}[\|\boldsymbol{\Lambda}_{t+1}\|_2]}{\gamma} \leqslant \sqrt{D_{\mathcal{X}} L_f} t^{1/2} \mathcal{O}(T^{1/4}) + \sqrt{\mathcal{D}(\mathcal{X}, \psi)}\mathcal{O}(T^{1/2}) + (C_\psi^2 + \frac{2\sigma^2}{K})^{1/2} t^{1/2} \mathcal{O}(1)$$

Now, we have:

$$\alpha \sum_{\tau=1}^{t} \mathbb{E}[\|\boldsymbol{\Lambda}_\tau\|_2] \leqslant \sqrt{\mathcal{D}(\mathcal{X},\psi)}t\mathcal{O}(T^{-1/2}) + \sqrt{D_{\mathcal{X}}L_f}\mathcal{O}(t^{3/2})\mathcal{O}(T^{-3/4})$$
$$+ (C_\psi^2 + \frac{2\sigma^2}{K})^{1/2}\mathcal{O}(t^{3/2})\mathcal{O}(T^{-1}).$$

Consequently:

$$\text{h-CFit}_t^r \leqslant \sqrt{\mathcal{D}(\mathcal{X},\psi)}\left[\mathcal{O}(\sqrt{T}) + t\mathcal{O}(T^{-1/2})\right] + \sqrt{D_{\mathcal{X}}L_f}\left[\sqrt{t}\mathcal{O}(T^{1/4}) + \mathcal{O}(t^{3/2})\mathcal{O}(T^{-3/4})\right]$$
$$+ (C_\psi^2 + \frac{2\sigma^2}{K})^{1/2}\left(\sqrt{t}\mathcal{O}(1) + \mathcal{O}(t^{3/2})\mathcal{O}(T^{-1})\right)$$

Since $t \leqslant T$, the result follows. ■

7.5. Discussions on the parameters and constants

This section aims to show the possibility of improving GOMSP by adapting the mirror map to the underlying feasible set. To this end, we compare the constants arising in the performance guarantees given in the previous section, both if the Euclidean norm -, and if the smoothed entropy serves as the regularizer. Throughout this section, we consider the constraint set:

$$\mathcal{X} = \left\{\boldsymbol{x} \in \mathbb{R}^D_{\geqslant 0} : \sum_{i=1}^{D} x^i \leqslant B\right\}, \quad B \geqslant 1$$

$\mathcal{D}(\mathcal{X},\psi)$ and $\mathcal{D}_{\mathcal{X}}^{\psi}$ To compute $\mathcal{D}(\mathcal{X},\psi_\epsilon^{\text{ent}})$, notice first that ψ_ϵ is strictly convex and therefore the minimizer of this function is an extreme point of $\mathcal{X}$. Consequently, we have for $\epsilon \leqslant 1$:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \psi_\epsilon^{\text{ent}}(x) = B\ln(B).$$

Now, by KKT-argumentations, it yields for $\epsilon \geqslant e^{-1}$:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \psi_\epsilon^{\text{ent}}(x) = -D\epsilon\ln(\epsilon).$$

Combining both observations, we have:

$$\mathcal{D}(\mathcal{X},\psi_\epsilon^{\text{ent}}) = B\ln(B) + D\epsilon\ln(\epsilon).$$

In contrast, we have:

$$\mathcal{D}(\mathcal{X},\|\cdot\|_2^2/2) = B^2.$$

So, using the smoothed entropy yields better dependency of $\mathcal{D}(\mathcal{X},\psi)$ on B ($B\ln(B)$ vs. B^2). However, $\mathcal{D}(\mathcal{X},\psi)$ has a linear dependency on the D which one fortunately can offset by choosing $\epsilon \in [e^{-1}, 1)$ large enough. The constant $\mathcal{D}_{\mathcal{X}}^{\psi}$ is irrelevant for our consideration, since it is equal B for both choices of ψ.

L_ψ and Sensitivity to Variation Elementary computation yields

$$L_{\psi_\epsilon^{\mathrm{ent}}} = \max\left\{|1 + \ln(\epsilon)|, |1 + \ln(B + \epsilon)|\right\}.$$

If $\epsilon \in [e^{-1}, 1]$, this quantity simplifies to:

$$L_{\psi_\epsilon^{\mathrm{ent}}} = 1 + \ln(B + \epsilon).$$

In contrast, we have:

$$L_{\|\cdot\|_2} = B.$$

We see that choosing $\psi = \psi_\epsilon^{\mathrm{ent}}$ instead of $\psi = \|\cdot\|_2^2/2$ might yields an improvement of the dependency of L_ψ on B ($\ln(B)$ vs. B) and therefore an improvement of the dependency of the GOMSP's regret performance on the variation. However, as a different choice of mirror map leads to a different norm measuring the variation, caution is required to this regard: With the choice $\psi = \psi_\epsilon^{\mathrm{ent}}$ we measure the variation by means of $\|\cdot\|_1$, and with the Euclidean norm as regularizer, we measure the variation by means of $\|\cdot\|_2$ which is in general smaller than $\|\cdot\|_1$ (by at worst the factor $\sqrt{D}$). The discussion in this paragraph is irrelevant for the h-CFit guarantee given in the previous section, since it is independent of the path variation.

Constants related to loss function and penalty function Clearly, $C_{3,\psi}$ is equal in both choices of the regularizer. Since $\|\cdot\|_\infty \leqslant \|\cdot\|_2$, $C_{1,\psi_\epsilon^{\mathrm{ent}}}$ might be smaller than $C_{1,\|\cdot\|_2^2/2}$. Similar argumentation yields that $C_{1,\psi_\epsilon^{\mathrm{ent}}}$ might be smaller $C_{1,\|\cdot\|_2^2/2}$.

Strong Convexity and Noise It is immediate to see that:

$$K_{\|\cdot\|_2} = 1.$$

Moreover, by Proposition 2.3, we have:

$$K_{\psi_\epsilon} = \frac{1}{B}.$$

So in case $B > 1$, GOMSP with $\psi_\epsilon^{\mathrm{ent}}$ as the regularizer might suffer more from noise amplification than GOMSP with the Euclidean norm as regularizer. Our advice concerning this issue is to normalize as far as possible the problem such that the restated problem has $B = 1$. Regarding the power σ_ψ of the persistent noise itself, we can leverage from choosing the smoothed entropy over the Euclidean norm as the regularizer of GOMSP. To see this, consider, for instance, an i.i.d. noise $(\boldsymbol{M}_t)_t$ where the coordinates of $\boldsymbol{M}_t$ are independent standard Gaussian random variables. It holds that $\sigma^2_{\|\cdot\|_2^2/2}$ is of order D. In contrast, $\sigma^2_{\psi_\epsilon^{\text{ent}}}$ is of order $\ln(D)$ which is better.

7.6. Numerical Simulation

In order to verify our theoretical findings, we test GOMSP and present in this section the result of our simulations. We first begin by stating the setting in our experiment.

7.6.1. Online Problem Setting

We test our method on a special case of the problem setting stated in Example 13 with 20 generators and 10 constraints, i.e.:

$$D = 20 \quad \text{and} \quad R = 10,$$

described in the following:

Feasible Set We consider the feasible set:

$$\mathcal{X} = \left\{ \boldsymbol{x} \in \mathbb{R}^D : \sum_{i=1}^{D} x_i \leqslant B \right\}, \quad \text{with } B = 1.$$

The reason for choosing $B = 1$ is to prevent possible noise amplification by using a regularizer other than the Euclidean norm (see paragraph *d*) in Section 7.5). For other setting where $B \neq 1$ one may reformulate the online problem such that the resulted feasible set has $B = 1$.

Loss Function We consider the quadratic cost function:

$$c_t^{(i)}(x^{(i)}) = \mathbf{a}_t^{(i)}(x^{(i)})^2 + \boldsymbol{b}_t^{(i)}\boldsymbol{x}^{(i)},$$

where:

$$a_t^{(i)} = 0.5\sin(\pi t/50) + 5 + \tilde{a}_t,$$

with $(\tilde{a}_t)$ is an i.i.d. random sequence uniformly distributed in the interval $[0, 0.5]$, and where $b_t^{(i)} = 0.5\sin(\pi t/100) + 6 + \tilde{b}_t$, where $(\tilde{b}_t)$ is an i.i.d. random sequence uniformly distributed in the interval $[0, 0.2]$. We set the demand service constant to be:

$$\boldsymbol{M} = 20.$$

Our model for the time-varying non-stationary demand is given as :

$$d_t = 0.1\cos(\pi t/125) + 0.7 + \tilde{d}_t,$$

where $(\tilde{d}_t)$ is an i.i.d. random sequence uniformly distributed in the interval $[0, 0.2]$.

Constraints The constraints are described by the quadratic functions

$$E_t^{i\to j}(\boldsymbol{x}^{(i)}) = c^{i\to j}(\boldsymbol{x}^{(i)})^2 + e^{i\to j}\boldsymbol{x}^{(i)},$$

where $c^{i\to j}$ and $e^{i\to j}$ are independent uniformly distributed random variable on the unit interval. We assume that the constraint thresholds are time-variant and non-stationary of the form:

$$E_t^{\max,j} = 0.05\cos(\pi t/50) + 0.2 + \tilde{e}_t,$$

where $(\tilde{d}_t)$ is an i.i.d. random sequence uniformly distributed in the interval $[0, 1]$.

7.6.2. Algorithm setting and Benchmarks

All the method which we apply to the online learning problem receives a warm start of the amount of 40 time-slots. Subsequently, we run the algorithms for $T = 500$. We test GOMSP on the online learning problem describe previously with both the smoothed entropy with $\epsilon = 0.5$ and Euclidean norm as a regularizer, where we set the step size to be $\gamma = 0.1/\sqrt{T}$ and the regularization parameter to be $\alpha = 15\gamma$. As choices of h we consider $h = [\cdot]_+$ and $h = [\cdot]_+^2$.

Noisy Feedback To model the disturbance of the gradient feedback, we assume that learner can only observe the cost coefficients $(\mathbf{a}_t^{(i)})_i$ and $(\boldsymbol{b}_t^{(i)})_i$ at time t up to a Gaussian random disturbance. Specifically, we assume at time t that the learner sees $(\hat{a}_t^i)_i$ and $(\hat{b}_t^i)_i$, where:

$$\hat{\mathbf{a}}_t^{(i)} = \mathbf{a}_t^{(i)} + \tilde{\boldsymbol{M}}_{t+1}^{(i),1} \quad \text{and} \quad \hat{b}_t^{(i)} = \mathbf{a}_t^{(i)} + \tilde{\boldsymbol{M}}_{t+1}^{(i),2},$$

with $(\tilde{\boldsymbol{M}}_t^{(i),1})_{i,t}$ (resp. $(\tilde{\boldsymbol{M}}_t^{(i),2})_{i,t}$) is the sequence of i.i.d. mean zero Gaussian random variable with standard deviation $\sigma_a > 0$ ($\sigma_b > 0$). Throughout our simulation, we set

$$\sigma_a = 0.2 \quad \text{and} \quad \sigma_b = 1.$$

MOSP We compare GOMSP with the modified online saddle-point (MOSP) introduced in [30] with fixed primal and dual step size equal to

$$\gamma \in \{0.1, 0.07, 0.05\}/\sqrt{T}.$$

In contrast to the works [30, 35], we simulate MOSP with imperfect gradient feedback with the noise structure described in the previous paragraph. we The (noisy) update of this algorithm is given by:

$$\boldsymbol{X}_{t+1} = \boldsymbol{\Pi}_{\mathcal{X}}\left[\boldsymbol{X}_t - \gamma\left(\hat{\mathbf{v}}_t + \sum_{r=1}^{N}\left[\nabla(\mathrm{h}\circ\mathbf{g}_t^r)(\boldsymbol{X}_t)\right]\boldsymbol{\Lambda}_t^{(r)}\right)\right]$$
$$\boldsymbol{\Lambda}_{t+1} = \boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}}\left[(1-\alpha\gamma)\boldsymbol{\Lambda}_t + \gamma\mathrm{h}(\mathbf{g}_t(\boldsymbol{X}_t))\right].$$

ODG Furthermore, we also compare GOMSP with the stochastic dual gradient (SDG) method (see e.g., [30, 137, 138]), which we modify as follows:

$$\boldsymbol{X}_{t+1} \in \arg\min_{\boldsymbol{x}\in\mathcal{X}} \hat{f}_t(x) + \langle\boldsymbol{\Lambda}_t, \mathbf{g}_t(x)\rangle$$
$$\boldsymbol{\Lambda}_{t+1} = [\boldsymbol{\Lambda}_t + \gamma\mathbf{g}_t(\boldsymbol{X}_t)]_+,$$

where $\hat{f}_t$ is the loss function with perturbed coefficients as described in Paragraph a). This modification is for the sake of fairness in the comparison since the original SDG method requires non-causal knowledge and does not consider the possibilities of disturbance in the feedback.

7.6.3. Simulation Result

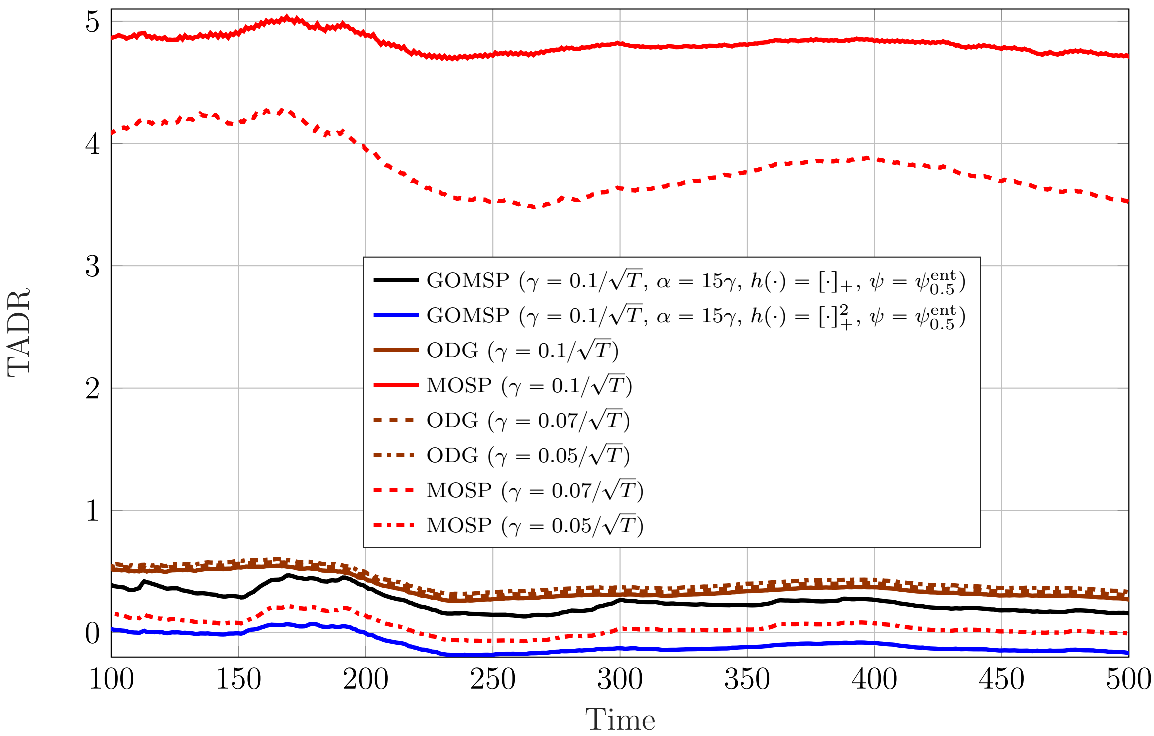


Figure 7.1.: Time-average Dynamic Regret (TADR) for GOSMP and benchmarks ODP and MOSP with perturbed cost $\sigma_a = 0.2$ and $\sigma_b = 1$.

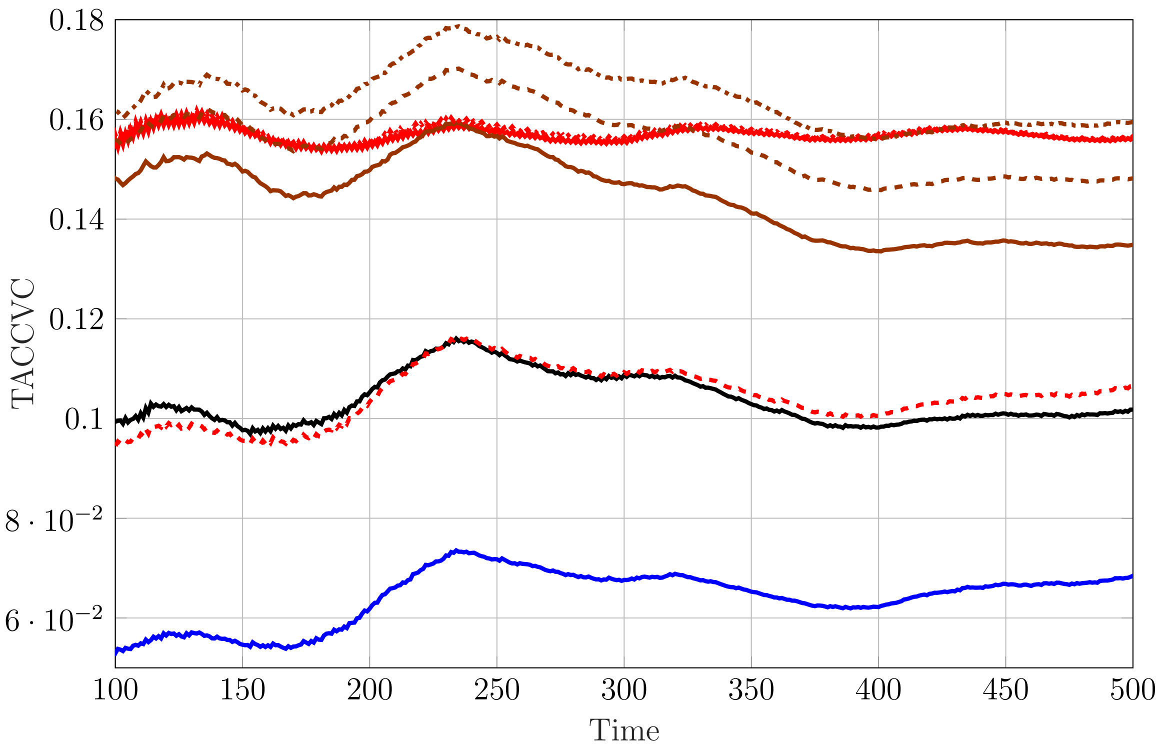


Figure 7.2.: Time-average clipped constraint violation (TACCV) for GOSMP and benchmarks ODP and MOSP with perturbed costs. For legend see Fig. 7.1.

Clipped Constraint Violation At first, we evaluate the Time average clipped constraint violation (TACCV) of the different methods given by:

$$\frac{\sum_{t=1}^{t}\sum_{r=1}^{R}[\mathbf{g}_t^r(\boldsymbol{X}_t)]_+}{tR}.$$

We see that in case the step sizes of the methods coincide ($\gamma = 0.1$), GOMSP with smoothed entropy as the regularizer, independent of the choice of h, clearly outperform ODG and MOSP. However, we see that $h = [\cdot]_+^2$ yields the best performance. Moreover, even by reducing the step sizes of ODG and MOSP to $\gamma = 0.07/\sqrt{T}$ and $\gamma = 0.05/\sqrt{T}$ the corresponding TACCV is still higher than that of GOMSP.

Dynamic Regret Now we examine the dynamic regret of the methods averaged over time (TADR). We provide the plot of this quantity in Fig. 7.1. Our method clearly outperform ODG w.r.t. to the performance measure TADR in the case where the step sizes of MOSP and its benchmarks coincide ($\gamma = 0.1/\sqrt{T}$). However, running MOSP with smaller step size ($\gamma = 0.05/\sqrt{T}$) it outperforms GOMSP with $h = [\cdot]_t$. This occurence can however be changed by choosing $h = [\cdot]_+^2$ since GOMSP possesses in this case the lowest and even negative dynamic regret.

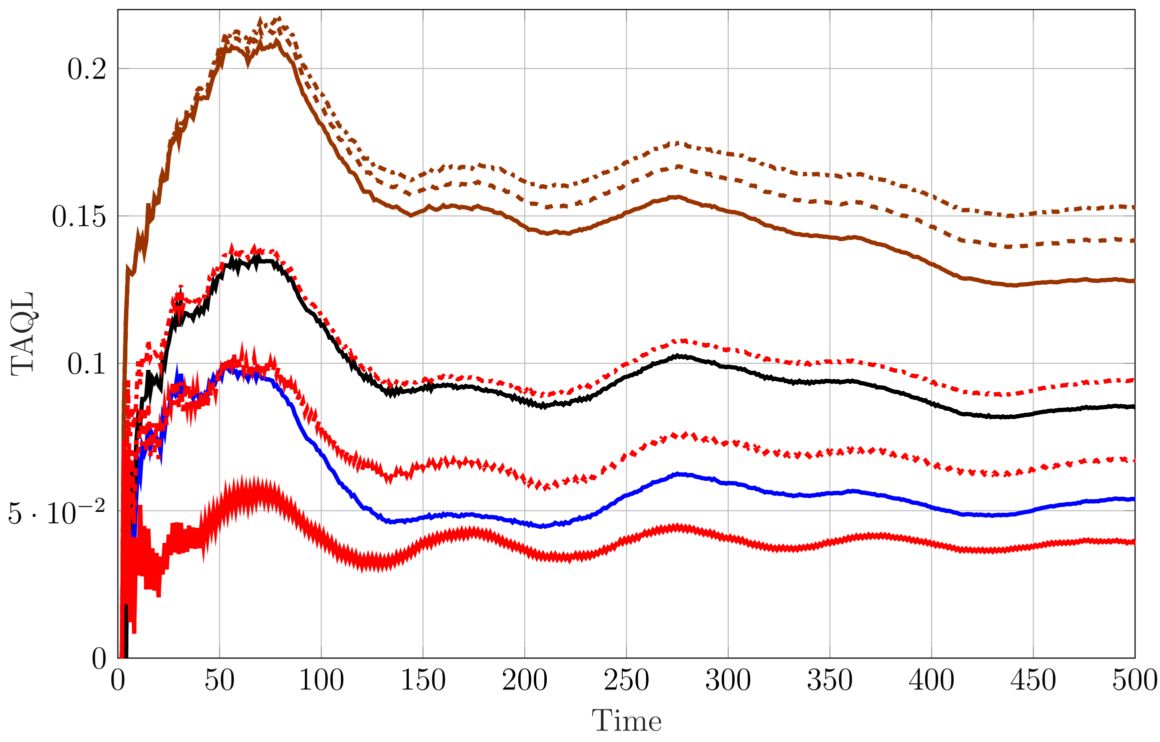


Figure 7.3.: Time-average Queue Length (TAQL) for GOSMP and benchmarks ODP and MOSP with perturbed cost $\sigma_a = 0.2$ and $\sigma_b = 1$. For legend see Fig. 7.1.

Queue Length In our experiment, we also examine the queue length $(Q_t)_t$ of GOMSP and its benchmarks, which is given by

$$Q_{t+1}^r = [Q_t^r + \mathbf{g}_t^r(\boldsymbol{X}_t)]_+,$$

with $Q_0^r = 0$. This quantity is relevant for applications where the current constraint violation can be compensated by previous actions that are strictly constraint fulfilling, which occurs in systems having the ability to buffer (see e.g. [30]). Clearly, small TACCV does not imply small queue length since the former implies that the constraint violations remain small and the latter allows some substantial constraint violations of cost constraint values strictly smaller than the allowed threshold. We plot the time-average queue length (TAQL) $\sum_{r=1}^R Q_t^r/tR$ in Figure 7.3. We see that MOSP with $\gamma = 0.1/\sqrt{T}$ yields the lowest queue length. However, by observing its trajectory, this performance is caused by the fact that the update of MOSP highly and rapidly oscillates between states which strictly fulfilling the constraint and states violating the constraints. Such a behavior is not tolerable in technical applications since it might incur an additional switching cost (see e.g. [139]. Furthermore it is surprising in the face of the previous discussion on the difference between TACCVC and TAQL that ignoring the MOSP with $\gamma = 0.1/\sqrt{T}$, it is possible that GOMSP may have the smallest TAQL.

Impact of Mirror Map Choice At last, we are interested in investigating to what extent does the choice of the mirror map impacts the performance of GOSMP. Toward this end,

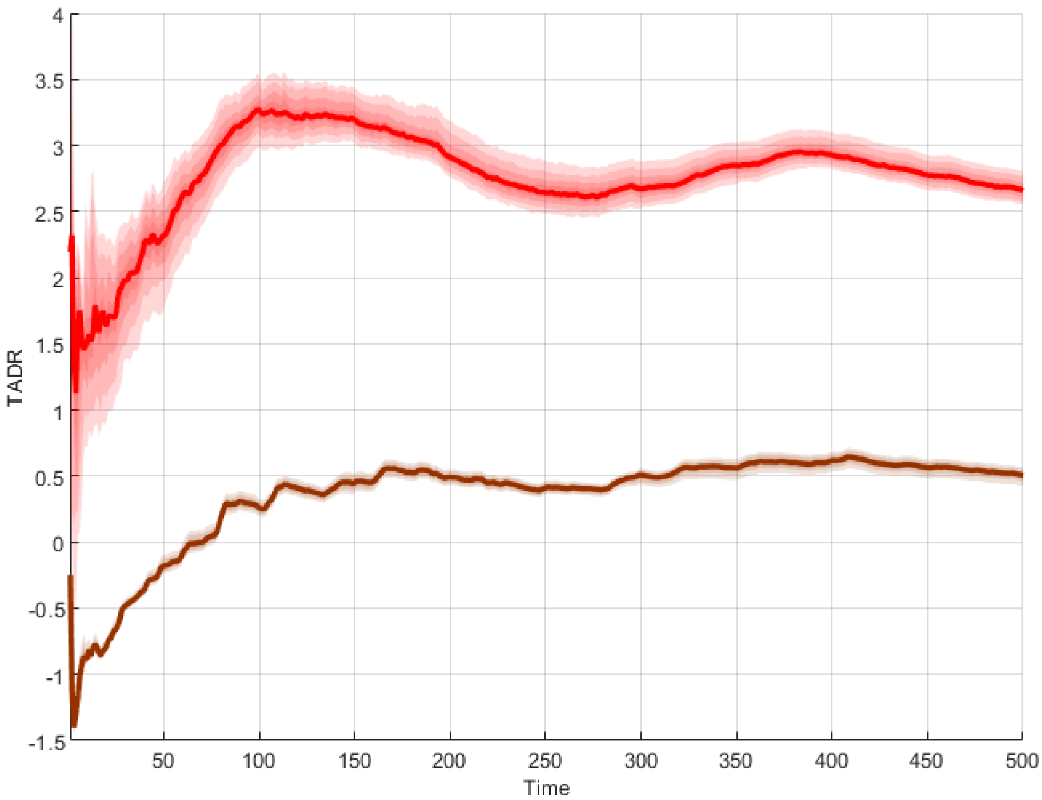


Figure 7.4.: TADR for GOSMP. Brown line corresponds to the sample average of TACCV in the smoothed entropy ($\epsilon = 0.5$) case and red line resp. in the Euclidean case. Shaded areas are each corresponds to 25%-, 50%-, 75%-, and 90%-percentile.

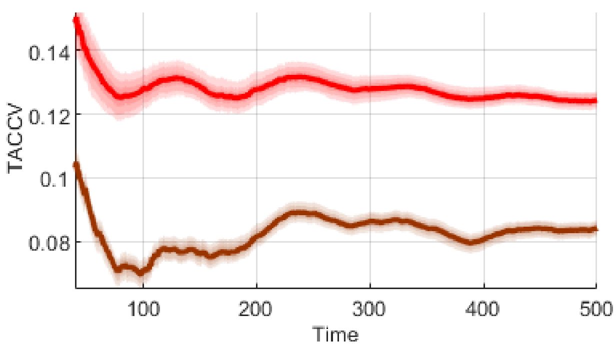


Figure 7.5.: TACCVC for GOSMP. Brown line corresponds to the sample average of TACCVC in the smoothed entropy ($\epsilon = 0.5$) case and red line resp. in the Euclidean case. Shaded areas are each corresponds to 25%-, 50%-, 75%-, and 90%-percentile.

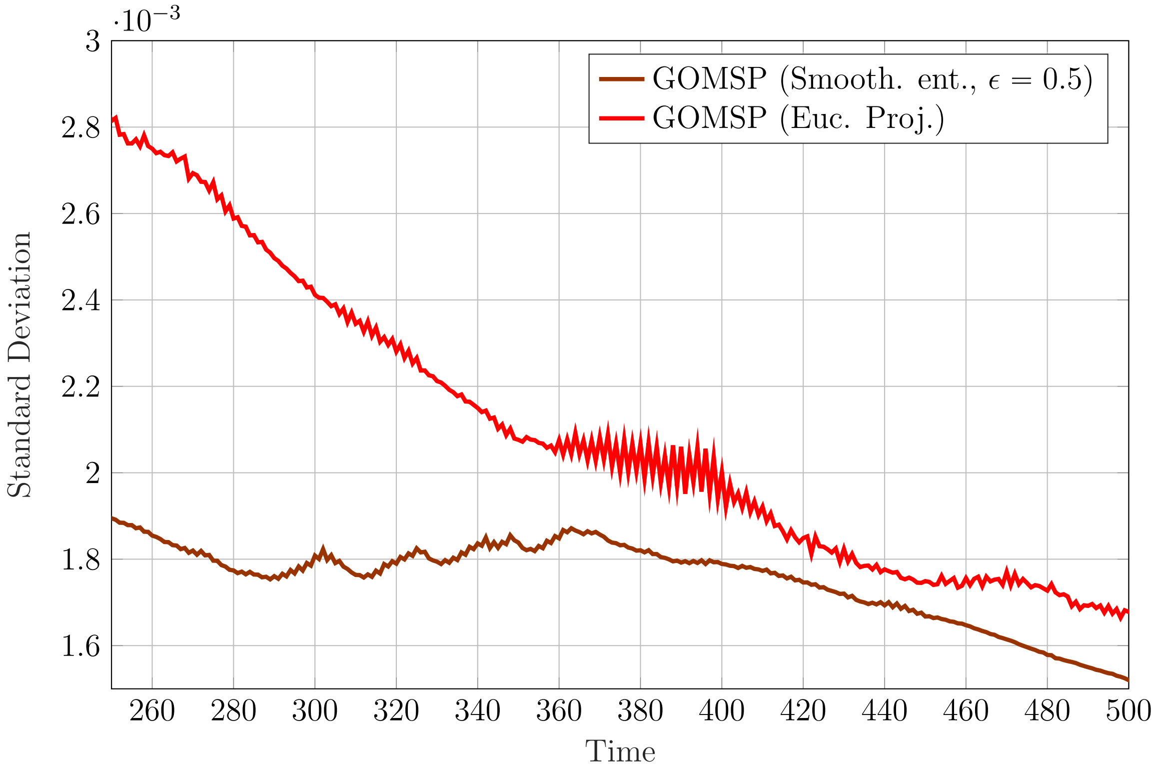


Figure 7.6.: Standard Deviation of GOSMP

we perform GOSMP with Euclidean projection and the smoothed entropy with $\epsilon = 0.5$ as regularizers. In both cases, we choose $\gamma = 0.1/\sqrt{T}$, $\alpha = 0.1/\sqrt{T}$, and $h = [\cdot]_+$. We simulate both instances of GOSMP with 200 gradient noise samples. Figure 7.4 depicts the dynamic regret of our simulation. There the thick line corresponds to the sample average of the trajectories, and the shaded line specifies the area where 25%, 50%, 75%, and 90% of the samples are. A clear trend which we can observe is that the TADR of GOSMP with smoothed entropy as regularizer is significantly lower than the TADR of GOSMP with Euclidean projection as regularizer. We believe that this effect aligns with the discussion made in paragraph c) in Section V. Moreover we observe that the TADR of GOSMP with smoothed entropy as regularizer is more volatile than that of GOSMP with Euclidean projection as regularizer. This observations confirms the hypothesis that using a mirror map other than the Euclidean one results in more robust algorithm behavior. We also observe similar trends in the resource-aware behavior of GOSMP (see Figure 7.5 and 7.6). However, the effect of noise reduction is less pronounced comparing to that of TADR.

7.7. Appendix

7.7.1. Missing Proofs in Section 7.4

The proof of Lemma 7.2 is straightforward following [25]:

Proof (Proof of Lemma 7.2): By inserting the primal iterate of the GOMSP into the bound given in Proposition 2.4, by using triangle inequality, by the inequality $(\sum_{i=1}^K a_i)^2 \leqslant K \sum_{i=1}^K a_i^2$, it is straightforward to obtain:

$$\begin{aligned}
&\Delta\mathcal{E}_t^1(x) \leqslant -\gamma\langle \boldsymbol{X}_t - \boldsymbol{x}, \nabla \mathrm{f}_t(\boldsymbol{X}_t) + [\nabla(\mathrm{h} \circ \mathbf{g}_t)(\boldsymbol{X}_t)]^T \boldsymbol{\Lambda}_t \rangle \\
&+ \gamma \tilde{\boldsymbol{M}}_{t+1} + \frac{1}{K}\left(\gamma^2 C_1^2 \|\boldsymbol{\Lambda}_k\|_2^2 + 2\gamma^2 (C_2^2 + \|\boldsymbol{M}_{t+1}\|_*^2)\right)
\end{aligned}$$

■

Our aim now is to proof Lemma 7.3. It is an immediate consequence of the following auxiliary statements:

Lemma 7.11: *It holds:*

$$\Delta\mathcal{E}_t^2 \leqslant \gamma\langle \boldsymbol{\Lambda}_t, \mathrm{h}(\mathbf{g}_t(\boldsymbol{X}_t))\rangle - (\alpha\gamma - \alpha^2\gamma^2)\|\boldsymbol{\Lambda}_t\|_2^2 + \gamma^2 C_3^2,$$

where $C_3 > 0$ *is a constant satisfying* (7.13).

Proof: It holds:

$$\begin{aligned}
&\|\boldsymbol{\Lambda}_{\tau+1}\|_2^2 = \|\boldsymbol{\Pi}_{\mathbb{R}^R_{\geqslant 0}}\left[(1-\alpha\gamma)\boldsymbol{\Lambda}_\tau + \gamma \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau))\right]\|_2^2 \\
&\leqslant \|\boldsymbol{\Lambda}_\tau + \gamma \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau)) - \alpha\gamma\boldsymbol{\Lambda}_\tau\|_2^2 = \|\boldsymbol{\Lambda}_\tau\|^2 \\
&+ 2\left[\gamma\langle \boldsymbol{\Lambda}_\tau, \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau))\rangle - \alpha\gamma\|\boldsymbol{\Lambda}_\tau\|^2\right] + \gamma^2\|\mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau)) - \alpha\boldsymbol{\Lambda}_\tau\|_2^2,
\end{aligned}$$

where the inequality follows from the usual property of the Euclidean projection operator. Triangle inequality, the inequality $(a+b)^2 \leqslant 2a+2b$, and (7.13) give $\|\mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau)) - \alpha\boldsymbol{\Lambda}_\tau\|_2^2 \leqslant 2\left(C_3^2 + \alpha^2\|\boldsymbol{\Lambda}_\tau\|_2^2\right)$. So combining all the derived inequalities, we obtain Lemma 7.3. ■

Lemma 7.12: *Suppose that h is monotone and g is convex. Let be τ fixed. It holds for any $\tilde{\boldsymbol{x}} \in \mathcal{Q}_\tau$:*

$$\langle \boldsymbol{\Lambda}_\tau, \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau))\rangle \leqslant \langle [\nabla(\mathrm{h} \circ \mathbf{g}_\tau)(\boldsymbol{X}_\tau)]^{\mathrm{T}} \boldsymbol{\Lambda}_\tau, \boldsymbol{X}_\tau - \tilde{\boldsymbol{x}}\rangle.$$

Proof: Let be $x \in \mathcal{X}$. Since h is monotone and $g^{(r)}$ is convex for all $r \in [R]$, it follows that $\mathrm{h} \circ g^{(r)}$ is convex. This and the fact that $\boldsymbol{\Lambda}_\tau \geqslant 0$ for all τ gives:

$$\begin{aligned}
\langle \boldsymbol{\Lambda}_\tau, \mathrm{h}(\mathbf{g}_\tau(\boldsymbol{X}_\tau))\rangle &\leqslant \langle \boldsymbol{\Lambda}_\tau, \mathrm{h}(\mathbf{g}_\tau(x))\rangle - \langle \boldsymbol{\Lambda}_\tau, \partial(\mathrm{h} \circ \mathbf{g}_\tau)(\boldsymbol{X}_\tau)(x - \boldsymbol{X}_\tau)\rangle \\
&= \langle \boldsymbol{\Lambda}_\tau, \mathrm{h}(\mathbf{g}_\tau(x))\rangle + \langle [\partial(\mathrm{h} \circ \mathbf{g}_\tau)(\boldsymbol{X}_\tau)]^{\mathrm{T}} \boldsymbol{\Lambda}_\tau, \boldsymbol{X}_\tau - \boldsymbol{x}\rangle
\end{aligned}$$

Since h is monotone, we have for $\tilde{\boldsymbol{x}} \in \mathcal{Q}_\tau \subset \mathcal{X}$, $\mathrm{h}(\mathbf{g}_\tau(\tilde{\boldsymbol{x}})) \leqslant 0$. Consequently since $\boldsymbol{\Lambda}_\tau \geqslant 0$, it yields $\langle \boldsymbol{\Lambda}_\tau, \mathrm{h}(\mathbf{g}_\tau(\tilde{\boldsymbol{x}}))\rangle \leqslant 0$. Combining all the computations, we obtain the desired result. ■

Consequently by combining Lemmas 7.11 and 7.12 we obtain Lemma 7.3.

Part III.

Distributed Coordination Algorithms

8. Stochastic Dynamic of First-Order Flocking-based Distributed Optimization

Abstract: We study a continuous first-order distributed gradient descent with flocking term of pure attracting force. We propose and analyze a stochastic variant in which the gradient is contaminated by Gaussian noise. We provide a bound of the distance between the objective value of the averaged iterate of each agents and the consensus optimum both in expectation and in probability. We discuss the interaction between the underlying apriori parameters (parameter of the functions and connectivity of the agents) of the problem, the parameters of the dynamic (step size/gradient weight and communication strength between the agents) and the volatility of the noise process.

8.1. Introduction

Problems of cooperative control in multi-agent systems have gained a lot of attention over the recent years due to emergence of large scale networks. Some examples in which such problem arised are smart grids [140], autonomous vehicle teams [141], processors in machine learning scenarios [142], sensor systems [143], multi-robot system [144] and cognitive networks [145, 146].

In many problems corresponding to such system, A global objective needs to be achieved by appropriate actions of agents. The difficulties arise by the fact that usually each agents can only access local information. One may canonically solve such problem in the centralized manner in the sense that for each time instance a central computing unit collects the processed local data from each agent and then subsequently gives command to all the agents based on it. However, this sort of system is sensitive to the failure of the central unit. Moreover, the communication between each agents and the central unit might be costly due to for instance the large distance between them. Another problem which could arise in the centralized system is that the computation done by the central unit might be infeasible for example due to the large number of participants in the system.

Numerous problem in the aforementioned applications is related to the problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^D} \frac{1}{N}\sum_{i=1}^{N} \mathrm{V}_i(\boldsymbol{x}). \tag{8.1}$$

where V_i is for each $i \in [N]$ a convex function only observable by the i-th agent. The specific task for each agent is to iterate the solution of (8.1) by means of the information about the observables and the iterate of other agents. First-order methods play an important or even crucial role in solving such sort of problem since higher-order information is often hard to obtain. For instance in large-scale machine learning applications this is due to the high dimension of the feature space and the overwhelming size of a typical data sets. Furthermore in many applications, one often has merely an inexact first-order oracle either desired due to e.g. computational infeasibility (see mini-batch method in machine learning), or measurement error, or adversarial attack. For instance in machine learning application, computing the gradient of a function is expensive since in most cases the corresponding function is the sum of large number of another functions. So instead of computing the full gradient one chooses randomly some functions in the sum and compute in that way an unbiased estimator of the full gradient. Another source of inexact random first-order oracle is for instance measurement error and adversarial attack.

A way to solve (8.1) is by the following algorithm where the iterates $X_{(i)}$ of i-th agent is specified by:

$$\begin{aligned} \mathrm{d}\boldsymbol{X}_{(i),t} &= \alpha_t \left[-\nabla \mathrm{V}_i(\boldsymbol{X}_{(i),t})\mathrm{d}t + \sigma_t \mathrm{d}B_t\right] && (8.2)\\ &\quad - \gamma_t \sum_{j\neq i}[A]_{i,j}(\boldsymbol{X}_{(i),t} - \boldsymbol{X}_{(j),t})\mathrm{d}t \ , && (8.3) \end{aligned}$$

where $(B_t)_{t\geqslant 0}$ denotes the Brownian motion. So for each time instance each agent call the gradient oracle perturbed by the Gaussian noise with power $\sigma_t \in \mathbb{R}_0^+$ for the locally observable function (see term (8.2)), collect the previous iterate of all neighboring agents, compare it with his past own iterate (see term (8.3)), and subsequently combined all the obtained information in order to update his own state. The algorithm (8.3) can be seen as stochastic gradient descent with flocking term of pure linear attractive force [147].

In this work we interpret the stochastic integral occuring in above description in the Itô's sense so that the continuous-time algorithm (8.3) can be seen as the continuous-time version of the discrete algorithm

$$\boldsymbol{X}_{(i),t} = \alpha_k[-\nabla \mathrm{V}_i(\boldsymbol{X}_{(i),k}) + \sigma_k \xi_k] - \gamma_k \sum_{j\neq i}[A]_{i,j}(\boldsymbol{X}_{(i),k} - \boldsymbol{X}_{(j),k}),$$

where $\xi_1, \xi_2, \ldots$ i.i.d. multivariate standard normal random variable. Thus the parameter α can in some sense be seen as the parameter determines the step size of the algorithm and γ as the parameter specifying the strength of communications between each agent.

Except in the applications where the continuous time method is admissible or even necessary such as in control problem, optimization methods are inherently discrete. However, continuous-time method is often easier and more convenient to handle since there are a huge literature on dynamical systems [148, 149] and control theory [150] providing useful techniques ([151, 152]). Moreover, it can provides heuristics for their design and analysis. An excellent example is the mirror descent algorithm which is motivated in continuous time [153]. Furthermore, more recent works [154–159] gives more intuitive view on Nesterov's accelerated first-order method by seeing the discretization of a second-order ordinary differential equations and based on that advices for the choices of the occuring parameters.

Our Contributions

In this paper, we analyze the continuous-time stochastic algorithm (8.3) with more general noise structure (see Remark 21). We investigate the role of the occuring parameters, viz. the step size, communication strength, and the volatility of the noise for the success of the algorithm. We quantify in Theorem 8.3 the consensus between the agents and provide in Theorem 8.5 a bound for the distance of the average function value of the averaged iterate of each agents to the optimum value in expectation under the convexity and boundedness of gradient assumption. In case that in addition the corresponding objective function is strongly convex, we are even able to quantify in Theorem 8.7 and Theorem 8.8 the probability of the aforementioned bound.

Relation to Prior Work

The first work on discrete gradient descent distributed algorithm can be traced back to the seminal works [160, 161]. Since then many extensions has been made. Among them, approaches which are closely related to our work is given e.g. in [162, 163]. In [164], the algorithm (8.3) without noise has been analyzed in a more general setting where the flocking term is substituted by the gradient of a certain convex function different from the objective. Under constant step size α, they derive conditions on the parameter β s.t. the algorithm converges. In contrast to our work, their statement is more of asymptotic nature.

Another alternative deterministic continuous algorithms which based on feedback control were provided and analyzed in e.g. [165, 166] for several cases of dynamic networks. However in contrast to our analysis, they rely to the strong convexity of the objective.

For the case where the summands of the objective function in (8.1) coincides, similar continuous stochastic algorithm as in (8.3) was already analyzed in [167, 168]. However their aim is different to ours since the flocking based algorithm they proposed serves rather as a tool to reduce the effect of noise modeled by the Brownian motion caused by the inexact first-order oracle.

To the best of our knowledge, there is until now no attempt to provide a non-asymptotic probability bound for the distance between iterate and the optimal point, since the works on (distributed) stochastic gradient descent either provide expectation bounds [167, 168] or long term asymptotic concentration proof [**?**]

8.2. Preliminaries

Spectral Graph and Consensus Subspace

$N \in \mathbb{N}$ denotes the number of agents. We model the information exchanging between the agents by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ without self-loop. The set of agents are represented by the set of vertices $\mathcal{V} = [N]$ and if agent i is able to interact with agent j then $\{i, j\} \in \mathcal{E}$. The adjacency matrix of $\mathcal{G}$ is denoted by $\mathbf{A}$ and is defined as the matrix for which $[\mathbf{A}]_{ij} = 1$ if $\{i, j\} \in \mathcal{E}$ and $[\mathbf{A}]_{ij} = 0$ else. The Laplacian $\mathbf{L}$ associated to $\mathcal{G}$ is defined as the matrix for which:

$$[\mathbf{L}]_{ij} = \begin{cases} \sum_k [\mathbf{A}]_{ik} & \text{if } i = j \\ -[\mathbf{A}]_{ij}, & \text{if } i \neq j. \end{cases}$$

This matrix is known to be symmetric and positive semidefinite and hence its real eigenvalues can be ordered in the non-decreasing way $0 \leqslant \lambda_1(\mathbf{L}) \leqslant \ldots \leqslant \lambda_N(\mathbf{L})$. It is known that $\lambda_1(\mathbf{L}) = 0$ with the eigenvector $\mathbf{1}_N$ whose elements are ones. We assume that the graph is connected which gives $\lambda_2(\mathbf{L}) > 0$. Moreover it holds $\lambda_2(\mathbf{L}) \leqslant N$ and inequality occurs in the case that the graph is complete. For a detailed treatment of those aspects, we refer to standard textbooks e.g. [169].

The space $\mathbb{R}^N \otimes \mathbb{R}^D$ denotes simply $\mathbb{R}^{ND}$ equipped with the scalar product:

$$\langle\!\langle \boldsymbol{x}, \boldsymbol{y} \rangle\!\rangle = \frac{1}{N} \sum_{i=1}^{N} \left\langle \boldsymbol{x}_{(i)}, \boldsymbol{y}_{(i)} \right\rangle,$$

and the induced norm:

$$|\!|\!| \boldsymbol{x} |\!|\!| = \langle\!\langle \boldsymbol{x}, \boldsymbol{x} \rangle\!\rangle.$$

Let be $i \in [N]$. The operation $(\cdot)_{(i)} : \mathbb{R}^N \otimes \mathbb{R}^D \to D$ used before is the operation of taking the i-th D-dimensional vector out of x, explicitly:

$$\boldsymbol{x}_{(i)} := (\boldsymbol{x})_{(i)} := (\boldsymbol{x}_{1+(i-1)D}, \ldots, \boldsymbol{x}_{iD}).$$

We define the consensus subspace as:

$$\mathcal{C} := \left\{ \mathbf{1}_N \otimes x : \quad x \in \mathbb{R}^D \right\}.$$

It is easy to see that the family

$$\{\mathbf{1}_N \otimes \mathbf{e}_{i,D}\}_{i=1}^{D},$$

constitute an orthonormal basis for $\mathcal{C}$, where $\mathbf{e}_{i,D}$, $i \in [N]$ denotes the canonical basis of $\mathbb{R}^D$. Let be $x \in \mathbb{R}^D \otimes \mathbb{R}^N$. The barycenter $\boldsymbol{x}^c$ of x is defined by the equation:

$$\mathbf{1}_N \otimes \boldsymbol{x}^c = \mathbf{P}_{\mathcal{C}} x,$$

where:

$$\mathbf{P}_{\mathcal{C}} := \mathbf{1}_N \mathbf{1}_N^T \otimes \mathbf{I}_D / N$$

denotes the orthogonal projection onto $\mathcal{C}$. $\boldsymbol{x}^c$ can explicitly be written as:

$$\boldsymbol{x}^c = \frac{\sum_{i=1}^N \boldsymbol{x}_{(i)}}{N}.$$

We define the fluctuation $\boldsymbol{x}^f$ between the vector $x \in \mathbb{R}^N \otimes \mathbb{R}^D$ and its barycenter $\boldsymbol{x}^c \in \mathbb{R}^D$ by:

$$\boldsymbol{x}^f := (\mathbf{I}_{\mathbb{R}^N \otimes \mathbb{R}^D} - \mathbf{P}_{\mathcal{C}})\, x = x - \mathbf{1}_N \otimes \boldsymbol{x}^c$$

Denote $\mathbf{I}_D \in \mathbb{R}^D \times \mathbb{R}^D$ the identity matrix. We often consider the "blown up" Laplacian $\mathbf{L} \otimes \mathbf{I}_D \in \mathbb{R}^{N \times N} \otimes \mathbb{R}^{D \times D} = \mathbb{R}^{ND} \times \mathbb{R}^{ND}$. This matrix is symmetric and its spectrum coincides with the spectrum of $\mathbf{L}$. Since the eigenspace corresponds to the zero eigenvalue is $\mathcal{C}$ it holds by the Courant-Fischer formula (see e.g. 4.2.11 Theorem in [170]) that for every $x \in \mathcal{C}^\perp$ with $|||\boldsymbol{x}||| = 1$:

$$\langle\langle \boldsymbol{x}, (\mathbf{L} \otimes \mathbf{I}_D)\, x \rangle\rangle \geqslant \lambda_2(\mathbf{L}) > 0. \tag{8.4}$$

Furthermore since $\mathcal{C}$ is the nullspace of $\mathbf{L} \otimes \mathbf{I}_D$ and $\mathbf{L} \otimes \mathbf{I}_D$ is symmetric, it holds by previous observation that for every $x \in \mathbb{R}^N \otimes \mathbb{R}^D$, $\langle\langle \boldsymbol{x}, (\mathbf{L} \otimes \mathbf{I}_N)\, x \rangle\rangle \geqslant \lambda_2(\mathbf{L}) |||\boldsymbol{x}^f|||^2$. We summarize those facts in the following Lemma:

Lemma 8.1: *Consider the mapping $\Phi : \mathbb{R}^N \otimes \mathbb{R}^D \to \mathbb{R}$ given by:*

$$\Phi(\boldsymbol{x}) = \frac{\langle x, (\mathbf{L} \otimes \mathbf{I}_N) x \rangle}{2}. \tag{8.5}$$

It holds:

1. *Φ is a bilinear non-negative bounded functional with $2\Phi(\boldsymbol{x}) \leqslant \lambda_N(\mathbf{L}) \|\boldsymbol{x}\|^2$.*
2. *For every $z \in \mathcal{C}$ and $x \in \mathbb{R}^N \otimes \mathbb{R}^D$, $\langle x, (\mathbf{L} \otimes \mathbf{I}_N)(\boldsymbol{x} - z) \rangle = 2\Phi(\boldsymbol{x} + z) = 2\Phi(\boldsymbol{x})$.*
3. *For every $x \in \mathcal{C}^\perp$, $\lambda_2(\mathbf{L}) \|\boldsymbol{x}\|^2 \leqslant 2\Phi(\boldsymbol{x})$.*

Proof: The fact that Φ is bilinear is obvious. The fact that $\Phi \geqslant 0$ follows from the fact that $\mathbf{L} \otimes \mathbf{I}_N$ is positive semidefinite. For the second property, notice that $\mathcal{C} = \mathcal{N}(\Phi)$.

Combining with the fact that $\mathbf{L}$ is symmetric and therefore also $\mathbf{L} \otimes \mathbf{I}_N$, we have:

$$\begin{aligned}\frac{\Phi(\boldsymbol{x}+z)}{2} &= \langle(\mathbf{L} \otimes \mathbf{I}_N)(\boldsymbol{x}+z), x+z\rangle = \langle(\mathbf{L} \otimes \mathbf{I}_N)x, x+z\rangle = \langle x, (\mathbf{L} \otimes \mathbf{I}_N)(\boldsymbol{x}+z)\rangle \\ &= \langle x, (\mathbf{L} \otimes \mathbf{I}_N)x\rangle = \frac{\Phi(\boldsymbol{x})}{2},\end{aligned}$$

where the The last property follows from (8.4). ■

Another easy but helpful observation is that for all $x \in \mathbb{R}^N \otimes \mathbb{R}^D$ and $y \in \mathbb{R}^D$:

$$\|\boldsymbol{x} - \mathbf{1} \otimes y\|^2_{N \otimes D} = \|\boldsymbol{x}^f\|^2_{N \otimes D} + \|\mathbf{1} \otimes \boldsymbol{x}^c - \mathbf{1} \otimes y\|^2_{N \otimes D} \tag{8.6}$$

since $\mathcal{C}^\perp \ni \boldsymbol{x}^f \perp (\mathbf{1} \otimes \boldsymbol{x}^c - \mathbf{1} \otimes y) \in \mathcal{C}$. Notice that since $\Phi \geqslant 0$ and $\Phi(\boldsymbol{x}) = \Phi(\boldsymbol{x} - \mathbf{1}_N \otimes \boldsymbol{x}^c)$ one may see Φ as a functional which measures the distance of a point $x \in \mathbb{R}^N \otimes \mathbb{R}^D$ to the consensus subspace.

Stochastic Analysis

In this work we assume throughout that filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with $\mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0}$ a filtration on $\mathcal{F}$, satisfying the usual condition. A stochastic process X on $\mathbb{R}_0^+ \times \Omega$ is called $\mathbb{F}$-adapted if $\boldsymbol{X}_t$ is $\mathcal{F}_t$-measureable for every $t \geqslant 0$. It is called continuous if it is almost surely continuous. It is a martingale if it is integrable (w.r.t. $\mathbb{P}$) and if $\mathbb{E}[\boldsymbol{X}_s|\mathcal{F}_t] = \boldsymbol{X}_t$ (a.s.) for all $s > t$.

We denote $(B_t)_{t \geqslant 0}$ a $\mathbb{F}$-adapted ND-dimensional Brownian motion (see e.g. Definition 4.3 in [171]). The integral w.r.t. the Brownian motion is understood in the Itô's sense. Let Y be a vector-valued process. We denote $\langle Y \rangle_t$ the covariation process matrix with $[\langle Y \rangle_t]_{i,j} = \langle Y_i, Y_j \rangle_t$ denoting the quadratic covariation/sharp bracket[1] between Y_i and Y_j. For detailed treatment on those aspects we refer to e.g. [171, 172].

8.3. Model Description

We define:

$$\mathrm{V}(\boldsymbol{x}) = \sum_{i=1}^N \mathrm{V}_i(\boldsymbol{x}_{(i)}), \quad x \in \mathbb{R}^N \otimes \mathbb{R}^D.$$

If V_i, $i \in [N]$ is convex, we have that V is as sum of convex functions also convex. The problem (8.1) can be written as an unconstrained problem on $\mathbb{R}^N \otimes \mathbb{R}^D$ as follows:

$$\min_{\boldsymbol{x} \in \mathbb{R}^N \otimes \mathbb{R}^D} \tilde{V}(\boldsymbol{x}) := \min_{\boldsymbol{x} \in \mathbb{R}^N} \mathrm{V}(\boldsymbol{x}) + \Phi(\boldsymbol{x}). \tag{8.7}$$

[1]The quadratic variation between two scalar processes X and Y is defined as the accumulation of the infinitesimal covariation between the increments of X and Y. For a formal definition see e.g. pp. 32 in [172]

The gradient flow corresponding to the problem (8.7) has the dynamic:

$$\dot{\boldsymbol{X}}(t) = -\nabla\tilde{V}(\boldsymbol{x}_t) = -\nabla \mathrm{V}(\boldsymbol{x}_t) - (\mathbf{L}\otimes\mathbf{I}_N)\boldsymbol{X}_t.$$

In this work, we consider more generally the dynamic:

$$\dot{X}(t) = \alpha_t\left[-\nabla \mathrm{V}(\boldsymbol{x}_t) - \beta_t(\mathbf{L}\otimes\mathbf{I}_N)\boldsymbol{X}_t\right], \tag{8.8}$$

If not otherwise stated, we assume the following:

Assumption 8.1: *The functions $\alpha : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ and $\beta : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ are continuously differentiable, and V has a Lipschitz gradient.*

Notice that Cauchy-Lipschitz Theorem asserts in this case that (8.8) has a unique solution.

Now, let be:

$$\sigma : \mathbb{R}_0^+ \times \mathbb{R}^N \otimes \mathbb{R}^D \mapsto \mathbb{R}^N \otimes \mathbb{R}^D \times \mathbb{R}^N \otimes \mathbb{R}^D.$$

We may give the stochastic version of (8.8) as follows:

$$\mathrm{d}\boldsymbol{X}_t = \alpha_t\left(-\nabla \mathrm{V}(\boldsymbol{X}_t)\mathrm{d}t + \sigma_t(\boldsymbol{X}_t)\mathrm{d}B_t\right) + \alpha_t\beta_t(\mathbf{L}\otimes\mathbf{I}_N)\boldsymbol{X}_t. \tag{8.9}$$

For the sake of simplicity, we assume that the initial point X_0 is deterministic. Throughout, we assume the following:

Assumption 8.2: *V has a Lipschitz continuous gradient and $\sigma : \mathbb{R}_0^+ \times \mathbb{R}^N \otimes \mathbb{R}^D \to \mathbb{R}^N \otimes \mathbb{R}^D \times \mathbb{R}^N \otimes \mathbb{R}^D$, $(t, x) \mapsto \sigma_t(\boldsymbol{x})$ is Lipschitz continuous in x and continuous in t.*

Under this condition, it holds (see e.g. Section 2.2 in [171]) that the solution $\boldsymbol{X}_t$ of (8.9) is a unique continuous $\mathbb{F}$-adapted process fulfilling:

$$\mathbb{E}\left[\int_0^t \|\boldsymbol{X}_s\|^2 \mathrm{d}s\right] < \infty, \quad \forall t \geqslant 0. \tag{8.10}$$

Furthermore we define the mapping

$$\Sigma = \sigma^*\sigma.$$

Remark 21: Aside from the fact that it is Gaussian, our noise model is quiet general since we allow σ to be matrix-valued and therefore also both correlation between the agents and correlation between the coordinates of each agents. Moreover, the volatility matrix may in our case depend on the iterate $\boldsymbol{X}_t$ of the agents.

8.4. Analysis of the Stochastic Dynamic

We begin first by estimating the stochastic dynamic of the kinetic energy $|||\boldsymbol{X}_t - \boldsymbol{x}_*|||^2$ where $\boldsymbol{x}_*$ is the optimizer of V. In doing this we use the notation:

$$\mathcal{E}(\boldsymbol{x}, \boldsymbol{y}) := |||\boldsymbol{x} - \boldsymbol{y}|||^2/2.$$

Lemma 8.2: *Suppose that V is strongly convex with parameter $K_V > 0$. Then for the solution $\boldsymbol{X}_t$ of (8.9) and for $\boldsymbol{x}_* = \arg\min V$, it holds:*

$$\begin{aligned}
\mathrm{d}\mathcal{E}(\boldsymbol{x}_t, \boldsymbol{x}_*) \leqslant &- \alpha_t \left[\frac{\mathrm{V}(\boldsymbol{x}_t) - \mathrm{V}(\boldsymbol{x}_*)}{N} + K_V \mathcal{E}(\boldsymbol{X}_t, \boldsymbol{x}_*)\right] \mathrm{d}t &(8.11)\\
&- \lambda_2(\mathbf{L})\alpha_t\beta_t \left|\left|\left| \boldsymbol{X}_t^f \right|\right|\right|^2 \mathrm{d}t &(8.12)\\
&+ \alpha_t \langle\langle \boldsymbol{X}_t - \boldsymbol{x}_*, \sigma_t(\boldsymbol{X}_t)\mathrm{d}B_t \rangle\rangle &(8.13)\\
&+ \frac{\alpha_t^2}{2N} \, tr\left[\Sigma_t(\boldsymbol{X}_t)\right] \mathrm{d}t, &(8.14)
\end{aligned}$$

If V is merely convex, then above inequality holds for all $x^* \in \arg\min V \cap \mathcal{C}$ with $K_V = 0$.

Proof: Since:

$$\nabla_{\boldsymbol{y}} \mathcal{E}(y, \boldsymbol{x}_*) = (y - \boldsymbol{x}_*)/N \quad \text{and} \quad \nabla^2_{\boldsymbol{y}} \mathcal{E}(y, \boldsymbol{x}_*) = \mathbf{I}/N$$

we have by Itô formula (see e.g. 3.6 Theorem in [172]):

$$\begin{aligned}
\mathrm{d}\mathcal{E}(\boldsymbol{X}_t, \boldsymbol{x}_*) &= \langle \nabla_{\boldsymbol{X}_t} \mathcal{E}(\boldsymbol{X}_t, x^*), \mathrm{d}\boldsymbol{X}_t \rangle + \frac{1}{2} \nabla^2_{\boldsymbol{X}_t} \mathcal{E}(\boldsymbol{X}_t, x^*) : \mathrm{d}\langle X \rangle_t\\
&= \langle \nabla_{\boldsymbol{X}_t} \mathcal{E}(\boldsymbol{X}_t, x^*), \mathrm{d}\boldsymbol{X}_t \rangle + \frac{1}{2} \sum_{i=1}^N \nabla^2_{\boldsymbol{X}_{(i),t}} \frac{\|\boldsymbol{X}_{(i),t} - x^*_{(i)}\|^2}{2} : \mathrm{d}\left\langle X_{(i)} \right\rangle_t\\
&= \langle\langle \boldsymbol{X}_t - \boldsymbol{x}_*, \mathrm{d}\boldsymbol{X}_t \rangle\rangle + \frac{1}{2N} \operatorname{tr}\left(\mathrm{d}\langle X \rangle_t\right)\\
&= -\langle\langle \boldsymbol{X}_t - \boldsymbol{x}_*, \alpha_t \left[\nabla \mathrm{V}(\boldsymbol{X}_t) + \beta_t (\mathbf{L} \otimes \mathbf{I}_N)\boldsymbol{X}_t\right] \rangle\rangle \mathrm{d}t\\
&\quad + \alpha_t \langle\langle \boldsymbol{X}_t - \boldsymbol{x}_*, \sigma(\boldsymbol{X}_t)\mathrm{d}B_t \rangle\rangle + \frac{\alpha_t^2}{2N} \operatorname{tr}\left(\Sigma(\boldsymbol{X}_t)\right) \mathrm{d}t.
\end{aligned}$$

Since $\boldsymbol{x}_* \in \mathcal{C}$, we have by Lemma 8.1:

$$\langle\langle \boldsymbol{X}_t - \boldsymbol{x}_*, (\mathbf{L} \otimes \mathbf{I}_N)\, \boldsymbol{X}_t \rangle\rangle \geqslant \lambda_2(\mathbf{L}) \left|\left|\left| \boldsymbol{X}_t^f \right|\right|\right|^2.$$

Moreover since V is strongly convex and $\boldsymbol{x}_* = \arg\min V$ we have:

$$-\langle \boldsymbol{X}_t - \boldsymbol{x}_*, \nabla \mathrm{V}(\boldsymbol{X}_t) \rangle \leqslant -\mathrm{V}(\boldsymbol{X}_t) - \mathrm{V}(\boldsymbol{x}_*) - \tfrac{K_V}{2} \|\boldsymbol{X}_t - \boldsymbol{x}_*\|^2.$$

Combining all above observations, we obtain the desired statement. ∎

The term (8.11) corresponds to the drift of the iterate towards the optimum which is due to gradient flow with respect to the potential V. The term (8.12) corresponds to the drift of the iterate towards the consensus space which is due to the flocking term. In the case that the objective is merely convex (8.12) and therefore the minimum is not unique, (8.12) asserts that the flocking term guides the algorithm to the consensus. The impact of noise is reflected in the two last terms: The Itô term (8.13) reflects the infinitesimal quadratic deviation accumulated by noise throughout the time, the term (8.14) which reflects the accumulation of the instantaneous noise over time.

Now, above Lemma implies a bound on the distance between the function value of the averaged iterate and the optimum value and the consensus of the averaged iterate of each agents:

Theorem 8.3: *Suppose that V is strongly convex with parameter $K_V > 0$, and that $\mathcal{C} \cap \arg\min V \neq \varnothing$. Moreover, suppose that:*

$$\sigma_{*,t}^2 := \sup_{\boldsymbol{x}} tr(\Sigma_t(\boldsymbol{x})) < \infty, \quad \forall t \geqslant 0. \tag{8.15}$$

Then for the solution $\boldsymbol{X}_t$ of (8.8) with constant β and for $\boldsymbol{x}_* = \arg\min V$ it holds:

$$\frac{\mathrm{V}(\overline{\boldsymbol{X}}_t^{\alpha}) - \mathrm{V}(\boldsymbol{x}_*)}{N} \leqslant \frac{\mathcal{E}_0(\boldsymbol{x}_*)}{\int_0^t \alpha_s \mathrm{d}s} + \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{2N\int_0^t \alpha_s \mathrm{d}s} + \frac{M_t - K_V \int_0^t \alpha_s \mathcal{E}(\boldsymbol{X}_s, \boldsymbol{x}_*)\mathrm{d}s}{\int_0^t \alpha_s \mathrm{d}s} \tag{8.16}$$

$$\tilde{\mathcal{E}}_t^{\alpha} \leqslant \frac{\mathcal{E}_0(\boldsymbol{x}_*)}{2\beta\lambda_2(\mathbf{L})\int_0^t \alpha_t \mathrm{d}t} + \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{4N\beta\lambda_2(\mathbf{L})\int_0^t \alpha_s \mathrm{d}s} + \frac{M_t - K_V \int_0^t \alpha_s \mathcal{E}(\boldsymbol{X}_s, \boldsymbol{x}_*)\mathrm{d}s}{2\beta\lambda_2(\mathbf{L})\int_0^t \alpha_t \mathrm{d}t}, \tag{8.17}$$

where:

$$\tilde{\mathcal{E}}_t^{\alpha} := \mathcal{E}(\overline{\boldsymbol{X}}_t^{\alpha}, \mathbf{1}_N \otimes \overline{\boldsymbol{X}}_t^{\alpha,c}), \quad \mathcal{E}_0(\boldsymbol{x}_*) := \mathcal{E}(X_0, \boldsymbol{x}_*),$$

and:

$$M_t := \int_0^t \alpha_s \langle\!\langle \boldsymbol{X}_s - \boldsymbol{x}_*, \sigma_s(\boldsymbol{X}_s)\mathrm{d}B_s \rangle\!\rangle$$

If V is merely convex, then above inequality holds for all $x^* \in \arg\min V \cap \mathcal{C}$ with $K_V = 0$.

Proof: Denote:

$$\mathcal{E}_t := \mathcal{E}(\boldsymbol{X}_t, \boldsymbol{x}_*).$$

Since $\boldsymbol{x}_* \in \mathcal{C}$, we have by Lemma 8.1 $\Phi \geqslant 0$. Thus by Lemma 8.2 and (8.15):

$$\mathrm{d}\mathcal{E}_t \leqslant -\alpha_t \left[\frac{\mathrm{V}(\boldsymbol{X}_t) - \mathrm{V}(\boldsymbol{x}_*)}{N} + K_V \mathcal{E}_t\right] \mathrm{d}t + \frac{\alpha_t^2 \sigma_{*,t}^2}{2N} \mathrm{d}t + \mathrm{d}M_t$$

By integrating, multiplying both sides by $1/\int_0^t \alpha_s \mathrm{d}s$, and by noticing that $\mathcal{E}_t \geqslant 0$ it yields:

$$\frac{\int_0^t \alpha_s \frac{(\mathrm{V}(\boldsymbol{X}_s)-\mathrm{V}(\boldsymbol{x}_*))}{N}\mathrm{d}s}{\int_0^t \alpha_s \mathrm{d}s} \leqslant \frac{\mathcal{E}(X_0, \boldsymbol{x}_*)}{\int_0^t \alpha_s \mathrm{d}s} + \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2}{2N\int_0^t \alpha_s \mathrm{d}s} + \frac{M_t}{\int_0^t \alpha_s \mathrm{d}s}.$$

Applying Fubini's Theorem and then Jensen's inequality to the left side we obtain (8.16). The proof of (8.17) is similar to before, the starting point is the following observation: Since $\boldsymbol{x}_* \in \arg\min V$ it holds $\mathrm{V}(\boldsymbol{X}_t) \geqslant \mathrm{V}(\boldsymbol{x}_*)$. Therefore by Lemma 8.2:

$$\begin{aligned} \mathrm{d}\mathcal{E}(\boldsymbol{X}_t, \boldsymbol{x}_*) \leqslant & -\alpha_t \left[K_V \mathcal{E}(\boldsymbol{X}_t, \boldsymbol{x}_*) + 2\lambda_2(\mathbf{L})\beta \mathcal{E}(\boldsymbol{X}_t, \mathbf{1}_N \otimes \boldsymbol{X}_t^c)\right] \\ & + \frac{\alpha_t^2 \sigma_{*,t}^2 \mathrm{d}t}{2N} + \mathrm{d}M_t. \end{aligned}$$ ■

In order to proceed, we need to handle which is the stochastic term M_t, which is done in the next two subsections.

Expectation Bound

Theorem 8.4: *Suppose that V is convex, and that $\mathcal{C} \cap \arg\min V \neq \varnothing$. Moreover, suppose that* (8.15) *holds. Then for the solution $\boldsymbol{X}_t$ of* (8.8)*, for all $x^* \in \mathcal{C} \cap \arg\min \mathrm{V}(\boldsymbol{x})$, and for all $k \in [N]$, it holds:*

$$\mathbb{E}\left[\frac{\mathrm{V}(\overline{\boldsymbol{X}}_t^{\alpha}) - \min V}{N}\right] \leqslant \frac{\hat{\mathcal{E}}_0}{\int_0^t \alpha_s \mathrm{d}s} + \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{2N\int_0^t \alpha_s \mathrm{d}s} \tag{8.18}$$

$$\mathbb{E}\left[\left\|\overline{X}_t^{\alpha,f}\right\|^2\right] \leqslant \frac{\hat{\mathcal{E}}_0}{\beta\lambda_2(\mathbf{L})\int_0^t \alpha_t \mathrm{d}t} + \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{2N\beta\lambda_2(\mathbf{L})\int_0^t \alpha_s \mathrm{d}s}, \tag{8.19}$$

where $\hat{\mathcal{E}}_0 = \min\{\mathcal{E}_0(\boldsymbol{x}) : x \in \arg\min V \cap \mathcal{C}\}$.

Proof: By (8.10) it follows that the Itô integral M_t is a martingale and therefore by the tower property of the conditional expectation $\mathbb{E}[M_t] = \mathbb{E}[M_0] = 0$. The final step consist of taking the expectation over (8.16) and (8.17). ■

Combining both estimates given above we obtain:

Theorem 8.5: *Suppose that the condition in Theorem 8.4 is fulfilled and Moreover, suppose that there exists $G > 0$ s.t.:*

$$\left\|\nabla \mathrm{V}(\mathbf{1}_N \otimes \overline{\boldsymbol{X}}_{(k),t}^{\alpha})\right\| \leqslant G, \quad \forall t \geqslant 0,\ k \in [N] \quad a.s. \tag{8.20}$$

It holds:

$$\mathbb{E}[\frac{\sum_{l=1}^{N} V_l(\overline{\boldsymbol{X}}^{\alpha}_{(k),t})}{N} - \min \frac{\sum_{l=1}^{N} V_l}{N}] \leqslant \frac{\hat{\mathcal{E}}_0}{\int_0^t \alpha_s \mathrm{d}s} + \frac{G(\sqrt{N}+1)\sqrt{\hat{\mathcal{E}}_0}}{\sqrt{2\lambda_2(\mathbf{L})\beta}\sqrt{\int_0^t \alpha_s \mathrm{d}t}}$$
$$+ \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{2N\int_0^t \alpha_s \mathrm{d}s} + \frac{G\frac{\sqrt{N}+1}{\sqrt{N}}}{\sqrt{\lambda_2(\mathbf{L})\beta}}\sqrt{\frac{\int_0^t \alpha_s^2\sigma_{*,s}^2 \mathrm{d}s}{\int_0^t \alpha_s \mathrm{d}s}}, \quad \forall k \in [N] \tag{8.21}$$

Proof: Therefore by (8.16), we have a bound for $\mathbb{E}[\mathrm{V}(\overline{\boldsymbol{X}}^{\alpha}_t) - \mathrm{V}(\boldsymbol{x}_*)]$. Thus it remains to estimate $\mathbb{E}[\mathrm{V}(\mathbf{1}_N \otimes \overline{\boldsymbol{X}}^{\alpha}_{(k),t}) - \mathrm{V}(\overline{\boldsymbol{X}}^{\alpha}_t)]$. To this end, we first compute:

$$\mathbb{E}\left[\frac{\mathrm{V}(\mathbf{1}_N \otimes \overline{\boldsymbol{X}}^{\alpha}_{(k),t}) - \mathrm{V}(\overline{\boldsymbol{X}}^{\alpha}_t)}{N}\right] \leqslant G\ \mathbb{E}\left[\|\mathbf{1}_N \otimes \overline{\boldsymbol{X}}^{\alpha}_{(k),t} - \overline{\boldsymbol{X}}^{\alpha}_t\|_{N\otimes D}\right]$$

The first inequality follows from the convexity of V and Cauchy-Schwarz inequality and the requirement (8.20). The last inequality follows by the estimate:

$$\left\|\mathbf{1}_N \otimes \overline{\boldsymbol{X}}^{\alpha}_{(k),t} - \overline{\boldsymbol{X}}^{\alpha}_t\right\| \leqslant (\sqrt{N}+1)\left\|\overline{\boldsymbol{X}}^{\alpha,f}_t\right\|$$

Combining this with (8.17) together with the martingale property of M_t, Jensen's inequality, and the inequality $\sqrt{a+b} \leqslant \sqrt{a} + \sqrt{b}$, for $a, b > 0$, we obtain the desired statement.■

Probability bound

In order to give a probability bound, we need first to the following statement:

Lemma 8.6: *Suppose that V is smooth and that:*

$$\tilde{\sigma}_{*,t} := \sup_{\boldsymbol{x}} \lambda_{max}\left(\Sigma_t(\boldsymbol{x})\right) < \infty$$

Let be $\beta, \delta \geqslant 0$ and $x \in \mathbb{R}^N \otimes \mathbb{R}^D$. With probability at least $1 - \exp(-\beta\delta)$, it holds:

$$\int_0^t \alpha_s \langle\langle \boldsymbol{X}_s - \boldsymbol{x}_*, \sigma_s(\boldsymbol{X}_s)\mathrm{d}B_s\rangle\rangle - \frac{\beta}{2N}\int_0^t \alpha_s^2 \tilde{\sigma}_{*,s}^2 \|\boldsymbol{X}_s - \boldsymbol{z}\|^2 \mathrm{d}s \leqslant \delta$$

Proof: It holds:

$$\mathrm{d}\langle M\rangle_t = \frac{\alpha_t^2}{N}\|\Sigma_t(\boldsymbol{X}_t)\left[\boldsymbol{X}_t - \boldsymbol{z}\right]\|^2 \mathrm{d}t.$$

Notice that M and $\langle M\rangle$ are continuous adapted processes (see Theorem 5.13 in [171] and 5.3 Definition and 5.8 Theorem in [172]). Therefore for all $n \in \mathbb{N}$, τ_n defined as the infimum of $t \geqslant 0$ satisfying $|M_t| + \langle M\rangle_t \geqslant n$ is a stopping time (c.f. Theorem 3.2 in [171]), i.e. $\{\tau_n \leqslant t\} \in \mathcal{F}_t$ for all $t \geqslant 0$. Moreover we have for every $t \geqslant 0$, $\tau_n \uparrow t$ almost surely.

Now, consider the Itô process:

$$Y_t^{(n)} = \beta \int_0^{t\wedge\tau_n} \langle\!\langle \alpha_t \sigma_t^{\mathrm{T}}(\boldsymbol{X}_t)\left[\boldsymbol{X}_s - \boldsymbol{z}\right], \mathrm{d}B_s\rangle\!\rangle - \frac{\beta^2}{2}\langle M\rangle_{t\wedge\tau_n}.$$

So, Itô formula (see e.g. 3.6 Theorem in [172]) asserts that:

$$\mathrm{d}\exp(Y_t^{(n)}) = \beta Y_t^{(n)} \langle\!\langle \tilde{V}_{t\wedge\tau_n}, \mathrm{d}B_{t\wedge\tau_n}\rangle\!\rangle.$$

Consequently since $Y^{(n)}$ and $\tilde{V}_{(\cdot)\wedge\tau_n}$ is bounded we have from above that $\exp(Y^{(n)})$ is a martingale. Therefore by Chebysheff inequality $\mathbb{P}\left(Y_t^{(n)} \geqslant \beta\delta\right) \leqslant \exp(-\beta\delta)$. Letting n goes to infinity, Fatou's Lemma asserts that with probability at least $1 - \exp(-\beta\delta)$ we have:

$$\int_0^t \langle\!\langle \alpha_t \sigma_t^{\mathrm{T}}(\boldsymbol{X}_t)\left[\boldsymbol{X}_s - \boldsymbol{z}\right], \mathrm{d}B_s\rangle\!\rangle - \frac{\beta}{2}\langle M\rangle_t \leqslant \delta$$

Finally, we obtain the desired statement by the elementary inequality:

$$\|\Sigma(\boldsymbol{X}_t)y\| \leqslant \sup_{\boldsymbol{x}} \lambda_{\max}(\Sigma(\boldsymbol{x}))\|\boldsymbol{y}\|$$

which asserts that:

$$\mathrm{d}\langle M\rangle_t \leqslant \frac{\alpha_t^2\tilde{\sigma}_{*,t}^2}{N}|\!|\!|\boldsymbol{X}_t - \boldsymbol{z}|\!|\!|^2 \mathrm{d}t. \qquad \blacksquare$$

Combining above result and Theorem 8.3 it yields:

Theorem 8.7: *Suppose that V is strongly convex and that V fulfills* (8.20)*. Suppose that β is constant. If*

$$K_V - \alpha_t\tilde{\sigma}_{*,t}^2 \geqslant 0, \tag{8.22}$$

we have with probability at least :

$$1 - \exp\left(-\delta N \int_0^t \alpha_s \mathrm{d}s\right)$$

it holds for all $k \in [N]$:

$$\frac{1}{N}\sum_{l=1}^N V_l(\overline{\boldsymbol{X}}_{(l),t}^{\alpha}) - \min V/N \leqslant \frac{\mathcal{E}_0(\boldsymbol{x}_*)}{\int_0^t \alpha_s \mathrm{d}s} + \frac{\int_0^t \alpha_s^2\sigma_{*,s}^2 \mathrm{d}s}{2N\int_0^t \alpha_s \mathrm{d}s} + \delta,$$

and:

$$\left|\!\left|\!\left|\overline{\boldsymbol{X}}_t^{\alpha,f}\right|\!\right|\!\right|^2 \leqslant \frac{\mathcal{E}_0(\boldsymbol{x}_*)}{\beta\lambda_2(\mathbf{L})\int_0^t \alpha_s \mathrm{d}s} + \frac{\int_0^t \alpha_s^2\sigma_{*,s}^2 \mathrm{d}s}{2N\beta\lambda_2(\mathbf{L})\int_0^t \alpha_s \mathrm{d}s} + \frac{\delta}{\beta\lambda_2(\mathbf{L})}, \tag{8.23}$$

Proof: From Lemma 8.6 and (8.22) it holds:

$$M_t - K_V\int_0^t \alpha_s \mathcal{E}(\boldsymbol{X}_s, \boldsymbol{x}_*)\mathrm{d}s \leqslant \int_0^t \alpha_s(\alpha_s\tilde{\sigma}_{*,s}^2 - K_V)\mathrm{d}s + \delta \leqslant \delta,$$

with probability at least:

$$1 - \exp(-2N\delta \int_0^t \alpha_s \mathrm{d}s)$$

By combining this observation and Theorem 8.3, we obtain (8.7) and (8.23). ∎

The following result follows from Theorem 8.7 by the similar computations as in the proof of Theorem 8.5:

Theorem 8.8: *Suppose that the condition in Theorem 8.7 and* (8.20) *is fulfilled. It holds with probability at least* $1 - \exp\left(-\delta^2 N \int_0^t \alpha_s \mathrm{d}s\right)$ *for all* $k \in [N]$*:*

$$\frac{1}{N}\sum_{l=1}^{N} V_l(\overline{\boldsymbol{X}}^{\alpha}_{(k),t}) - \min V/N \leqslant \frac{\mathcal{E}_0(\boldsymbol{x}_*)}{\int_0^t \alpha_s \mathrm{d}s} + \frac{G(\sqrt{N}+1)}{\sqrt{\beta\lambda_2(\mathbf{L})}} \frac{\sqrt{\mathcal{E}_0(\boldsymbol{x}_*)}}{\sqrt{\int_0^t \alpha_s \mathrm{d}s}}$$
$$+ \frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{2N\int_0^t \alpha_s \mathrm{d}s} + \frac{G\frac{\sqrt{N}+1}{\sqrt{N}}}{\sqrt{\beta\lambda_2(\mathbf{L})}} \sqrt{\frac{\int_0^t \alpha_s^2 \sigma_{*,s}^2 \mathrm{d}s}{2\int_0^t \alpha_s \mathrm{d}s}} + \delta\left[\frac{G(\sqrt{N}+1)}{\sqrt{\beta\lambda_2(\mathbf{L})}} + \delta\right]$$

8.5. Case Study: Persistent and Vanishing Noise

Suppose each agents has the same time-variant volatility $\sigma : \mathbb{R}_0^+ \to \mathbb{R}_0^+$. In case that σ is constant, the bound in Theorem 8.5 turns to:

$$\mathbb{E}\Big[\frac{\sum_{l=1}^N V_l(\overline{\boldsymbol{X}}^{\alpha}_{(k),t})}{N} - \min \frac{\sum_{l=1}^N V_l}{N}\Big] \leqslant \frac{\hat{\mathcal{E}}_0}{\int_0^t \alpha_s \mathrm{d}s} + \frac{C\sqrt{\hat{\mathcal{E}}_0}}{\sqrt{2}\sqrt{\int_0^t \alpha_s \mathrm{d}s}}$$
$$+ \frac{D\sigma^2 \int_0^t \alpha_s^2 \mathrm{d}s}{2N\int_0^t \alpha_s \mathrm{d}s} + \frac{C\sqrt{D}\sigma\sqrt{\int_0^t \alpha_s^2 \mathrm{d}s}}{\sqrt{\int_0^t \alpha_s \mathrm{d}s}}, \tag{8.24}$$

where $C = G(\sqrt{N}+1)/\sqrt{\beta\lambda_2(\mathbf{L})}$. By appropriate choice of $\beta > 0$, this constant can be made arbitrarily small. In doing that, the connectivity of the network which is reflected in the quantity $\lambda_2(\mathbf{L})$ provides help. Moreover, the bound increases at worst linearly with D and if $\alpha_t = \mathcal{O}(1/t)$, we obtain convergence in expectation of order $\mathcal{O}(1/\sqrt{\log(t)})$. If the function is strongly convex with $K_V \geqslant \alpha_t\sigma^2$ and if $\delta > 0$ small enough, we have with probability $1-\mathcal{O}(1/t^{\delta N})$, that the bound (8.24) holds in probability with additional factor of order $\mathcal{O}(\delta)$. In case that the noise is vanishing, one may obtain a better bound. For instance if $\mathcal{O}(1/t)$ and for constant step size, we have the decay in expectation of order $\mathcal{O}(1/\sqrt{t})$ and the corresponding bound hold with additional factor of order $\mathcal{O}(\delta)$ with high probability $1 - \mathcal{O}(\exp(-\delta N t))$.

9. Mesoscopic Stability of the Friedkin-Johnsen Opinion Dynamics

Abstract: In this chapter we consider the setting of communicating agents, which are additionally subject to extrinsic influence in form of informational bias, e.g., fake news. We study the impact of extrinsic influence in the form of informational bias, to the population's opinion. Our main aim is to formally show that informational bias results in the establishment of mesoscopic stability, meaning that the population's opinion is cluster-dispersive. Toward this direction, we propose the novel notion of substochastic complementation, which provides an efficient way to approximate the population's dynamic by cluster dynamics. Motivated by this notion, we propose a novel measure for cluster-dispersion of opinion dynamic in the face of the informational bias and analyze it for several limit cases of disturbances by informational bias.

9.1. Introduction

Over the last of several decades, the study of networks has been an indispensable part of many research disciplines, such as biology, physics, and social sciences. Recently, the network-based view within the context of multi-agent systems has gained importance in engineering research due to the emerging complex architectures and interconnectedness in human-made systems, e.g., in IoT and sensor networks. Other factors which foster the utilization of network-based view in engineering are the availability of big data, the growing presence of the internet, and the increasing demand for decentralization of systems. The latter factor is founded by the fact that a decentralized system offers several potential advantages over a centralized one, such as the improvement of security of the system, the reduction of the expense of the necessary communication infrastructure, robustness of the system concerning individual agents' failure, and computational superiority in terms of the speed and the problem size. Also, the research field of signal processing has taken advantage of network-based perspective giving new insight and fostering the development of novel tools (see e.g., [173–175], and the references therein).

Opinion Dynamic Model - DeGroot Model One of the central tasks in the study of networks is to characterize the opinion dynamic in a network. The basic model of opinion dynamic is the so-called DeGroot model [176]. According to the DeGroot model, each agent updates her opinion by taking a convex combination of neighboring opinions.The central result in the literature of the DeGroot model is that the opinion dynamic's outcome is related to the underlying network statistics. Specifically, the population's opinion asymptotically reaches a consensus equal to the average of the population's member respective to the normalized eigenvector corresponding to the leading eigenvalue of the network's communication matrix. Thereby, the second leading eigenvalue of the network's matrix determines the convergence (see e.g. [176]). Although the DeGroot model is quite simple, it is widely used: For instance, in the social sciences, empirical studies (e.g., [177–181]) show that the DeGroot model can capture the real-world opinion dynamics. In the engineering literature, the DeGroot model serves as a method (see e.g., [182,183]), also known as the gossiping algorithm, to coordinate a multi-agent system to a consensus state. Exemplary signal processing applications of gossiping algorithms are distributed estimation, source localization, and compression. For a more detailed discussion on this aspect, we refer to e.g. [184], wherein also concise discussion on implementation issues is given, such as the amount of energy consumed in the network for gossiping, effects of quantization, and noise. In recent years, gossiping algorithms have been extended in several directions. One such extension is given in [185], presenting a framework for the design and the analysis of randomized gossip algorithms for average consensus in arbitrary connected networks, where pairs of nodes are chosen randomly to exchange the data. Further investigations of such randomized dynamic is given in [186–188]. Another extension worth to be mentioned is the extension considering agents in the network forming their opinion in the Bayesian manner [189–191]. Also, the DeGroot model serves serves as the basic of many distributed consensus algorithms, see e.g., [192–196].

Informational Bias in the Opinion Dynamic – Friedkin-Johnsen Model One of the drawbacks of the DeGroot model is that it only considers interpersonal influences between the agents in the form of informational exchange. However, in practice, agents are additionally subject to informational biases. An example of such an agent is the stubborn agent clinging to specific fixed opinions, e.g., her initial opinion, or, more generally: past opinions. Another example of informational bias is extrinsically given fake news. An extension of the DeGroot model aligned with the aspects above is the Friedkin-Johnsen model [197]. This model extends the DeGroot model by adding the possibility that the agent's opinion is built by additionally taking the average of the pooled opinions and another specific fixed opinions. This averaging effect tends to dampen the opinion mixing. Indeed a hint of this effect is the fact that the Friedkin-Johnsen dynamic converges generally no longer to a consensus. Instead, the limit is dictated by the given informational bias [197]. The opinion limit can further be characterized by the hitting probabilities of the canonical random walk on the underlying network and the voltage of the electrical network induced by the underlying graph (see [198]). Furthermore, the Friedkin-Johnsen model has an interesting interpretation in term of game theory, i.e., it can be seen as the result of a repeated best-response game where the agents/players aiming to minimize both, the discrepancy with the neighbors' opinion and the discrepancy with the intrinsic belief in the form of informational bias (see e.g. [199]).

Problem Formulation In this work, we aim to give a contribution to the answer to the following general question:

Question 1: How do the informational biases effects the opinion dynamic of a communicating population of agents?

Our particular concern is on the inter-agent communication dampening effect of informational biases preventing the mixing of agents' opinions and, consequently, the diversity of the population's opinion. Our emphasis is on answering the following question:

Question 2: How can we quantify the impact of informational bias and communication structure on the diversity of the opinion dynamic?

In answering this question, we choose the path of finding an effective method to uncouple the opinion dynamic into smaller cluster dynamics, each running parallelly and independently. In this direction, the problem we aim to solve is:

Question 3: How to approximate the opinion dynamic of the whole population by intracluster opinion dynamic?

By providing such an approximation, we can provide the answer to Question 2 by defining the approximation error as a measure of the inter-cluster diversity of the populations' opinion.

Furthermore, such an approximation would provide another advantage; i.e., it allows one to ease the handle tasks with large scale networks, which is a typical structure of modern applications since working with the whole network is in genral computationally intractable. An example of such tasks is predicting the outcome of Friedkin-Johnsen opinion dynamic, requiring, if done with the whole large-scale network, the inverse of a large matrix. Another example of such tasks is the placement of stubborn agents in order, e.g., to determine the opinion of the networked population, or to ensure certain properties such as polarization. This problem arises not only in social sciences but also in engineering , e.g., within the context of the containment problem in mobile networks. [200]. Also, with such an approximation, analysis related to the statistics of the networks would be relieved.

Contributions In this work we introduce the notion of the substochastic complementation of a matrix which is a generalization of the notion of the so-called stochastic complementation introduced in [201]. This allows us to provide an optimal approximation of the opinion dynamic of informationally biased by intracluster opinion dynamic. Based on the substochastic complementation, we provide a measure for clusterness of the opinion dynamic due to both the clusterness of the communication structure and the informational bias of the agents' opinion. As our final result, we provide an asymptotic analysis of the degree of clusterness of the opinion of certain population's cluster

Relation to Prior Works The notion of the substochastic complementation of a stochastic matrix is generalization of the notion of the stochastic complementation introduced in [201]. The aforementioned work concerns with the problem of uncoupling an irreducible Markov chain into several smaller independent chains. Moreover, the notion of stochastic complementation allows a unified, clear, and simple description of the Simon-Ando theory for nearly uncoupled systems developed in the seminal work [202].

9.2. Model description

9.2.1. Basic Notations and Notions

Some computations made in this work are based on the following well-known block matrix inversion formula:

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \tilde{\mathbf{B}}^{-1} & -\tilde{\mathbf{B}}^{-1}\mathbf{B}_{12}\mathbf{B}_{22}^{-1} \\ -\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\tilde{\mathbf{B}}^{-1} & \mathbf{B}_{22}^{-1} + \mathbf{B}_{22}^{-1}\mathbf{B}_{21}\tilde{\mathbf{B}}^{-1}\mathbf{B}_{12}\mathbf{B}_{22}^{-1} \end{bmatrix},$$

where:

$$\tilde{\mathbf{B}} := \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21},$$

and $\mathbf{B}_{11}$, $\mathbf{B}_{12}$, $\mathbf{B}_{21}$, and $\mathbf{B}_{22}$ are matrices such that the expressions given above make sense. In this work, we also make use the so-called Woodbury formula which is:

$$(\mathbf{A}_1 + \mathbf{A}_2\mathbf{A}_3\mathbf{A}_4)^{-1} = \mathbf{A}_1^{-1} - \mathbf{A}_1^{-1}\mathbf{A}_2(\mathbf{A}_3^{-1} + \mathbf{A}_4\mathbf{A}_1^{-1}\mathbf{A}_2)^{-1}\mathbf{A}_4\mathbf{A}_1^{-1},$$

for matrices $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4$ for which above expressions make sense. Let $\mathbf{A} \in \mathbb{R}^{M\times N}$ and $\mathcal{I}_1 \subseteq [M]$ and $\mathcal{I}_1 \subseteq [N]$. We denote the matrix resulting by simultaneously eliminating the row with indices $\mathcal{I}_1$ and the column with indices $\mathcal{I}_2$ by $\mathbf{A}_{\mathcal{I}_1\mathcal{I}_2}$. Suppose that $M = N$ and $\mathcal{I}_1 = \mathcal{I}_2 = \mathcal{I} \subseteq [N]$. We denote $\mathbf{A}_{\mathcal{I}_1\mathcal{I}_2}$ simply by $\mathbf{A}_{\mathcal{I}}$

Let $\mathbf{A} \in \mathbb{R}^{D_1\times D_2}$ be a non-negative matrix. We say $\mathbf{A}$ (row-)stochastic if it holds $\mathbf{A}\mathbf{1}_{D_2} = \mathbf{1}_{D_1}$. If $\mathbf{A}\mathbf{1}_{D_2} \leqslant \mathbf{1}_{D_1}$, then we say that $\mathbf{A}$ is substochastic. In case that there exists an entry of the substochastic matrix $\mathbf{A}\mathbf{1}_{D_2}$ strictly smaller than 1, then $\mathbf{A}$ is called a proper substochastic matrix.

Let $\mathbf{A} \in \mathbb{R}^{D\times D}$. We denote the spectral radius of $\mathbf{A}$ by $\rho(\mathbf{A})$. We say $\mathbf{A}$ is Schur-stable if the absolute value of all its eigenvalues is smaller than 1 or equivalently $\rho(\mathbf{A}) < 1$. $\mathbf{A}$ is said to be irreducible if for any i, j there exists $m \in \mathbb{N}$ s.t. $[\mathbf{A}^m]_{ij} > 0$

In this chapter, we often work with a directed graph, which is defined as the pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. A $\mathcal{G}$ is said to be strongly connected if there is a path from each vertex in the graph to every other vertex. To any matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ with non-negative entries, we can associate a graph $\mathcal{G}[\mathbf{A}] = ([N], \mathcal{E}[\mathbf{A}])$ by setting $(i,j) \in \mathcal{E}[\mathbf{A}]$ if $[\mathbf{A}]_{ij} > 0$. We say The matrix $\mathbf{A}$ is said to be adapted to graph $\mathcal{G}$ if $\mathcal{G}[\mathbf{A}] \subseteq \mathcal{G}$. Let be $i \in [N]$. We denote the set of neighbors of i by:

$$\mathcal{N}(i) := \{j \in \mathcal{V} : \ (i,j) \in \mathcal{E}\}.$$

The concept of irreducibility can also be related to the concept of connectedness of a graph: A matrix $\mathbf{A}$ is irreducible if and only if $\mathcal{G}[\mathbf{A}]$ is strongly connected.

9.2.2. Opinion Model: DeGroot Model

Let be $N \in \mathbb{N}$. We consider in this work a finite set of communicating agents $[N] = \{1, \ldots, N\}$. The basic model model for opinion spreading in a network is the so-called DeGroot opinion formation model. According to DeGroot model, the opinion of an agent at each time step is formed by taking the average of her neighbors' opinions. Formally, the DeGroot model considers the opinion dynamic specified as follows:

$$\boldsymbol{X}_i(k) = \sum_{j=1}^{N} a_{ij} \boldsymbol{X}_j(k-1), \tag{9.1}$$

where $\boldsymbol{X}_j(k)$ denotes the opinion of agent $j \in [N]$ at time $k \in \mathbb{N}_0$, and $a_{ij} \in \mathbb{R}_{\geqslant 0}$ represent the trust agent i puts in agent j. To ensure the averaging aspect one assumes that:

$$\sum_j a_{ij} = 1. \tag{9.2}$$

In order to express (9.1) in a compact way, one can define the so-called trust/communication matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ with:

$$[\mathbf{A}]_{ij} = a_{ij}, \quad \forall i,j \in [N].$$

This gives the following compact alternative formulation of (9.1):

$$\boldsymbol{X}(k) = \mathbf{A}\boldsymbol{X}(k-1),$$

where $\boldsymbol{X}(k)$ denotes the vector containing the opinion of all agents at time k. Align with the assumption (9.2), we require that the trust matrix $\mathbf{A}$ is a stochastic matrix.

The assumption that $\mathbf{A}$ is a stochastic matrix relates the study of DeGroot opinion dynamic to the study of markov chains and gives several interesting insight into the mixing process of the opinions of the communicating agents. One of the central result toward this direction is that under certain conditions on the trust matrix, the opinion of the agents coincides and equal to the average of the agents' initial opinion weighted by the mode of $\mathbf{A}$ corresponding to the largest eigenvalue of $\mathbf{A}$. We provide the specific statement in the following (see also [176]):

Proposition 9.1: *Suppose that $\mathcal{G}[\mathbf{A}]$ is strongly connected, then the dynamic (9.3) satisfies:*

$$\lim_{n\to\infty} \boldsymbol{X}_i(n) = \beta \mathbf{1}_D, \quad \textit{for an } \beta \in \mathbb{R},$$

if and only if $\mathbf{A}$ is aperiodic. Moreover:

$$\beta = \langle \pi, \boldsymbol{X}(0) \rangle,$$

where $\pi \in \Delta([N])$ satisfies $\pi^{\mathrm{T}}\mathbf{A} = \pi^{\mathrm{T}}$.

In this case: Consensus specified by the Perron-Frobenius eigenvector!

9.2.3. Opinion Model: Friedkin-Johnson Model

A straightforward way to extend the DeGroot opinion model is by introducing opinion bias in the opinions' update as follows:

$$\boldsymbol{X}_i(k) = \gamma_i \sum_{j=1}^{N} a_{ij} \boldsymbol{X}_j(k-1) + \lambda_i \xi_i(k), \tag{9.3}$$

where $\xi_i(k) \in \mathbb{R}$, $\gamma_i \in [0,1]$, and $\lambda_i \in [0,1]$. $\xi(k)$ models the opinion bias at time k preventing the the opinion mixing between the agents. The constant γ_i specify to what extent agent i is susceptible to her neighbors' opinion, and the constant λ_i to what extent the agent i clings to the informational bias. By this reason, we refer γ_i as the susceptibility constant and λ_i as the bias constant. By averaging reason, we set:

$$\lambda_i = 1 - \gamma_i.$$

An agent having the susceptability constant equal to zero, or equivalently, the bias constant equal to one, is insensitive against the opinion of other agent. We refer to this sort of agent as stubborn agent.

An example of informational bias is the initial opinion of the agent: Setting $\xi_i(k) = \boldsymbol{X}_i(0)$, we model agent i as an individual attaching to her initial opinion. Also, $\xi_i(k)$ can stand for extrinsic information, such as fake news, or also be set as $\xi_i(k) = \boldsymbol{X}_i(k-1)$, in order to model the agent i as a stubborn agent clinging to the previous opinion.

Using the notation of the trust matrix $\mathbf{A}$ specified in the previous subsection and defining the diagonal matrix $\mathbf{\Gamma} \in \mathbb{R}^{N\times N}$ having the susceptibility constants as entries:

$$\mathbf{\Gamma} = \operatorname{diag}(\gamma_1, \ldots, \gamma_N),$$

and the diagonal matrix $\mathbf{\Lambda} \in \mathbb{R}^{N\times N}$ having the bias constants as entries:

$$\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_N),$$

we can write the Friedkin-Johnsen dynamic (9.3) more compactly as:

$$\boldsymbol{X}(k) = \mathbf{\Gamma A}\boldsymbol{X}(k-1) + \mathbf{\Lambda}\xi(k) = \mathbf{\Gamma A}\boldsymbol{X}(k-1) + (\mathbf{I} - \mathbf{\Gamma})\xi(k). \tag{9.4}$$

Aiming to make the essential factors determining the agents' opinion transparent, we can iterate (9.4) and obtain:

$$\boldsymbol{X}(k) = (\mathbf{\Gamma A})^k \boldsymbol{X}(0) + \sum_{l=1}^{k} (\mathbf{\Gamma A})^{k-l} (\mathbf{I}_N - \mathbf{\Gamma})\xi(l). \tag{9.5}$$

So, according to above equation, it follows that the opinion of the agents depends on the initial opinion and the accumulation of the informational biases.

To further discuss the iterate (9.5) such as determining the asymptotic limit behavious of (9.5), it is advantageous to consider a more general class of dynamics than (9.5). For this sake, we define the class of affine dynamic by introducing the mapping:

$$\mathrm{St} : \mathbb{R}_{\geqslant 0}^{N\times N} \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^{\mathbb{N}_0},$$

with $\mathrm{St}[\mathbf{B}, u, x_0](0) = x_0$, and:

$$\mathrm{St}[\mathbf{B}, \boldsymbol{u}, \boldsymbol{x}_0](k+1) = \mathbf{B}\,\mathrm{St}[\mathbf{B}, \boldsymbol{u}, \boldsymbol{x}_0](k) + \boldsymbol{u}. \tag{9.6}$$

where $\mathbf{B} \in \mathbb{R}_{\geqslant 0}^{N\times N}$. We refer (9.6) as the affine dynamic induced by the Matrix $\mathbf{B}$ and the vector $\boldsymbol{u}$ with initial state $\boldsymbol{x}_0$. Now by iterating (9.6), it holds:

$$\mathrm{St}[\mathbf{B}, \boldsymbol{u}, \boldsymbol{x}_0](k) = \mathbf{B}^k \boldsymbol{x}_0 + \sum_{l=1}^{k} \mathbf{B}^{k-l}\boldsymbol{u} = \mathbf{B}^k \boldsymbol{x}_0 + \sum_{l=0}^{k-1} \mathbf{B}^{l}\boldsymbol{u}. \tag{9.7}$$

This identity gives rise to the following characterization of the asymptotic limit of (9.6):

Proposition 9.2: *Suppose that* $\mathbf{B}$ *is schur-stable. Then for any* u *and* x_0*, it holds:*

$$\lim_{k\to\infty} \mathrm{St}[\mathbf{B}, u, x_0](k) = (\mathbf{I} - \mathbf{B})^{-1} u$$

Proof: The fact that $\mathbf{B}$ is Schur-stable asserts that

$$\rho(\mathbf{B}) < 1. \tag{9.8}$$

By the reason that the spectral radius is submultiplicative, we have $\rho(\mathbf{B}^k) \leqslant \rho(\mathbf{B})^k$. Combining both observations, we have that

$$(\mathbf{B})^k \to 0 \quad \text{as} \quad k \to \infty. \tag{9.9}$$

Now, (9.8) gives that $(\mathbf{I}-\mathbf{B})$ is invertible, and therefore we have:

$$\sum_{l=1}^{k}\mathbf{B}^{k-l}=\sum_{l=0}^{k-1}\mathbf{B}^{k}\xrightarrow{k\to\infty}(\mathbf{I}-\mathbf{B})^{-1}. \tag{9.10}$$

Applying (9.9) and (9.10) into the formula (9.7), we have as desired:

$$\lim_{k\to\infty}\mathrm{St}[\mathbf{B},u,x_0](k)=\left(\lim_{k\to\infty}\mathbf{B}^k x_0\right)+\left(\lim_{k\to\infty}\sum_{l=1}^{k}\mathbf{B}^{k-l}u\right)=(\mathbf{I}-\mathbf{B})^{-1}u \qquad \blacksquare$$

Assuming that the weighted trust matrix $\mathbf{\Gamma A}$ is Schur-stable, we can provide by means of the above proposition the asymptotic limit of the Friedkin-Johnsen dynamic as follows:

$$\lim_{k\to\infty}\boldsymbol{X}(k)=(\mathbf{I}-\mathbf{\Gamma A})^{-1}(\mathbf{I}-\mathbf{\Gamma})\xi. \tag{9.11}$$

One observation from above equation is that with increasing time, the influence of the initial opinion to the population's opinion receeds, while the informational bias becomes the determining factor of the population's opinion. We can even specify the latter observation by the following simple Lemma:

Lemma 9.3: *Let $\mathbf{B}_1$ and $\mathbf{B}_2$ be square matrices with non-negative entries and suppose that $\mathbf{B}_1\mathbf{B}_2$ is Schur-stable. Then:*

1. *If $\mathbf{B}_2$ is stochastic, then $(\mathbf{I}-\mathbf{B}_1\mathbf{B}_2)^{-1}(\mathbf{I}-\mathbf{B}_1)$ is stochastic*

2. *If $\mathbf{B}_2$ is (resp. proper) substochastic, then $(\mathbf{I}-\mathbf{B}_1\mathbf{B}_2)^{-1}(\mathbf{I}-\mathbf{B}_1)$ is (resp. proper) substochastic.*

Proof: We only show that $\mathbf{B}_2$ is substochastic implies that $(\mathbf{I}-\mathbf{B}_1\mathbf{B}_2)^{-1}(\mathbf{I}-\mathbf{B}_1)$ is substochastic. The remaining statements follows by similar arguments. Since $\mathbf{B}_2$ is substochastic and $\mathbf{B}_1$ is non-negative, we have $\mathbf{B}_1\mathbf{B}_2\mathbf{1}\leqslant\mathbf{B}_1\mathbf{1}$, and thus:

$$(\mathbf{I}-\mathbf{B}_1\mathbf{B}_2)\mathbf{1}\geqslant(\mathbf{I}-\mathbf{B}_1)\mathbf{1}.$$

Since $(\mathbf{I}-\mathbf{B}_1\mathbf{B}_2)^{-1}$ is non-negative, it follows by multiplying both sides of above inequality by this term:

$$\mathbf{1}\geqslant(\mathbf{I}-\mathbf{B}_1\mathbf{B}_2)^{-1}(\mathbf{I}-\mathbf{B}_1)\mathbf{1}.$$

Thus the desired statement follows. $\blacksquare$

The fact that $(\mathbf{I}-\mathbf{\Gamma A})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ is stochastic, which is a consequence of above lemma, asserts that the opinions of the agents each approach the average of the informational biases ξ.

9.2.4. On Schur-Stability of Friedkin-Johnsen Dynamics

One assumption needed in order to analyze the Friedkin-Johnsen Dynamic for large times is the Schur-stability of the matrix $\mathbf{\Gamma A}$, ensuring the limit (9.11). The aim of this subsection is to discuss some properties of $\mathbf{\Gamma}$ and $\mathbf{A}$ leading to the truth of the aforementioned desired assumption.

One property which ensures the schur-stability of a matrix is the following:

Lemma 9.4: *Let* $\mathbf{B}$ *be a* $N \times N$ *irreducible sub-stochastic matrix having at least one row whose sum is strictly smaller than one. Then it is Schur-stable.*

Proof: By the Perron-Frobenius Theorem for irreducible non-negative matrix, we can find a strictly positive vector v summing to 1 satisfying:

$$v\mathbf{B} = \rho(\mathbf{B})v.$$

Now let us define the row vector ϵ having entries:

$$\epsilon_i = (1 - \sum_j [\mathbf{B}]_{ij})/N \geqslant 0,$$

and correspondingly the stochastic matrix:

$$\hat{\mathbf{B}} = \mathbf{B} + \epsilon \mathbf{1}^{\mathrm{T}}.$$

Thus we have:

$$\begin{aligned}\rho(\mathbf{B}) = \rho(\mathbf{B})\|v\| = \|\rho(\mathbf{B})v\|_1 = \|v\mathbf{B}\|_1 &= \sum_j \sum_i v_i [\mathbf{B}]_{ij} = \sum_j \sum_i v_i([\mathbf{B}]_{ij} + \epsilon_i - \epsilon_i) \\ &= \sum_j \sum_i ([\mathbf{B}]_{ij} + \epsilon_i)v_i - \sum_j \sum_i \epsilon_i v_i = \|\hat{\mathbf{B}}v\|_1 - N\langle v, \epsilon\rangle = 1 - N\langle v, \epsilon\rangle < 1,\end{aligned}$$

where the last inequality follows from the fact that $\mathbf{B}$ has at least one row having sum strictly smaller than one and therefore $\epsilon_i > 0$ for an i, and the fact that v have strictly positive entries. ∎

From above Lemma, we have that the matrix $\mathbf{\Gamma A}$ specifying the communication structure is Schur-stable if it is irreducible and has a row with the sum strictly smaller than 1. Intuitively, the latter condition means that at least an agent is subject to informational bias, i.e. $\lambda_i > 0$ or equivalently $\gamma_i < 1$ for an $i \in [N]$. The latter condition can be assumed to be satisfied, since otherwise, there is no-need to work with the Friedkin-Johnsen dynamic.

Now, in case that the agents are all not fully biased, or equivalently the agents still communicate with each other, the irreducibility of $\mathbf{\Gamma A}$ is fully characterized by the irreducibility of the communication matrix $\mathbf{A}$. Specifically, we have the following statement:

Lemma 9.5: *Suppose that $\gamma_i > 0$ for all $i \in [N]$. Then $\mathbf{\Gamma A}$ is irreducible if and only if $\mathbf{A}$ is irreducible.*

$$\gamma_i > 0, \quad \forall i \in [N].$$

Proof: If $\mathbf{\Gamma} > 0$ then $\mathcal{G}(\mathbf{\Gamma A}) = \mathcal{G}(\mathbf{A})$. So $\mathcal{G}(\mathbf{\Gamma A})$ is strongly connected if and only if $\mathcal{G}(\mathbf{A})$ is strongly connected. Since a matrix is irreducible if and only if its induced graph is strongly connected, we have the desired result ■

So from above Lemma we have in particular that if all agents communicate with each other and at least one agent is subject to informational bias, then $\mathbf{\Gamma A}$ is Schur-stable.

Now, suppose that not necessarily $\mathbf{\Gamma} > 0$. We have from Lemma 9.4 that $\mathbf{\Gamma A}$ is not irreducible. Therefore, 9.4 is not applicable to provide a further sufficient condition on $\mathbf{A}$ leading to the Schur-stability of $\mathbf{\Gamma A}$. So the question is how to ensure the latter property in face of the possibility of the existence of a stubborn agent. Toward this direction, we have the following statement:

Lemma 9.6: *Suppose that there exists $i \in [N]$ s.t. $\gamma_i = 0$, and denote set of indices $i \in [N]$ for which $\gamma_i = 0$ by $\mathcal{S}$. Then if $\mathbf{A}_{\mathcal{S}^c}$ is irreducible and if:*

$$\text{there exists } i \in \mathcal{S}^c \text{ and } j \in \mathcal{S} \text{ s.t.:} \quad [\mathbf{A}]_{ij} > 0, \tag{9.12}$$

then $\mathbf{\Gamma A}$ is Schur-stable.

Proof: By simultaneous row and column permutations we can transform $\mathbf{\Gamma A}$ into the following matrix block upper triangular matrix:

$$\mathbf{B} := \begin{bmatrix} \mathbf{\Gamma}_{\mathcal{S}^c}\mathbf{A}_{\mathcal{S}^c} & \tilde{\mathbf{B}} \\ 0 & 0 \end{bmatrix},$$

where $\tilde{\mathbf{B}}$ is a certain matrix. Consequently it holds: $\rho(\mathbf{\Gamma A}) = \rho(\mathbf{\Gamma}_{\mathcal{S}^c}\mathbf{A}_{\mathcal{S}^c})$. Now, by (9.12), we have that at least a row of $\mathbf{\Gamma}_{\mathcal{S}^c}\mathbf{A}_{\mathcal{S}^c}$ has the sum strictly smaller than 1. Consequently we have by 9.4, $\rho(\mathbf{\Gamma}_{\mathcal{I}^c}\mathbf{A}_{\mathcal{I}^c}) < 1$, as desired. ■

So, above Lemma asserts that in case stubborn agents exist in the system, as long as the partial communication matrix of the non-stubborn agent is irreducible, and as long as there is a non-stubborn agent influenced by the opinion of a stubborn agent, then the corresponding Friedkin-Johnsen dynamic has a limit (9.11).

From previous discussions, we see that irreducibility might be a quiet cumbersome assumption for the Schur-stability of $\mathbf{\Gamma A}$. In the following we provide a weaker condition for the ensurance of Schur-stability of $\mathbf{\Gamma A}$ known in the literature:

Lemma 9.7: *Let $\mathbf{B}$ be a proper substochastic matrix. Suppose that any node in $\mathcal{G}[\mathbf{B}]$ has a path to a deficiency node i in $\mathcal{G}[\mathbf{B}]$, i.e. a node i satisfying $\sum_j [\mathbf{B}]_{ij} < 1$. Then $\mathbf{B}$ is Schur-stable.*

Proof: The stated condition asserts that there exists $k_0 \in \mathbb{N}$ s.t. the rows of $\mathbf{B}^{k_0}$ is strictly smaller then 1. Consequently for any $n \in \mathbb{N}$ and $r \in \mathbb{N}$, we have:

$$\mathbf{B}^{k_0+r}\mathbf{1} = \mathbf{B}^{k_0}\mathbf{B}^{r}\mathbf{1} \leqslant \mathbf{B}^{k_0}\mathbf{1} < \mathbf{1}.$$

so for large enough n, we have $\mathbf{B}^n\mathbf{1}$ and consequently $\|\mathbf{B}^n\|_\infty^{1/n} < 1$. Finally Gelfand's formula asserts that:

$$\rho(\mathbf{B}) = \lim_{n\to\infty} \|\mathbf{B}^n\|_\infty^{1/n} < 1,$$

as desired. ■

So, from above Lemma, it follows that in order that the Friedkin-Johnsen dynamic is stable, it is enough all agents receive (indirectly and irrespective of finite time) the opinion of an agent subject to informational bias.

Remark 22: Clearly, above Lemma implies Lemma 9.6 and Lemma 9.5. The reason we state both lemmas earlier is since the literature often make use of irreducibility assumption.

9.2.5. Applications of Friedkin-Johnsen Dynamics

In order to give a motivation for working with (9.6) we provide some applications of some affine dynamics in the following:

Example 18 (PageRank Computation): In this application, the object of the study is a network consisting of web pages. We represent this network by graph, where the set of vertices correspond to the web pages and the edges correspond the links between the pages.Specifically, for two web pages $(i, j) \in \mathcal{E}$, we set $(i, j) \in \mathcal{E}$, if page i has an outgoing link to page j, or equivalently, if page j has an incoming link from page i. The aim of of the PageRank algorithm is to provide a measure of relevance of each web page $i \in \mathcal{C}$ by assigning each page a value $\boldsymbol{x}_i^* \in [0, 1]$, define as the solution of the equations:

$$\mathbf{M}\boldsymbol{x}^* = \boldsymbol{x}^* \quad \text{and} \quad \sum_{i=1}^{n} \boldsymbol{x}_i = 1,$$

where:

$$\mathbf{M} = (1-\tau)\mathbf{A} + \frac{\tau}{|\mathcal{V}|}\mathbf{1}\mathbf{1}^{\mathrm{T}},$$

with $\tau \in (0, 1)$ is a chosen constant, and $\mathbf{A}$ is defined as:

$$\mathbf{A}_{ij} := \begin{cases} \frac{1}{|\mathcal{N}(j)|} & \text{if } k \in \mathcal{N}(i) \\ 0 & \text{otherwise} \end{cases}.$$

One can under certain condition show that the PageRank vector can be computed by the distributed iterate:

$$\boldsymbol{x}(n+1) = \mathbf{M}\boldsymbol{x}(n) = (1-\tau)\mathbf{A}\boldsymbol{x}(n) + \frac{\tau}{|\mathcal{V}|}\mathbf{1}\mathbf{1}^{\mathrm{T}}\boldsymbol{x}(n).$$

Provided that $\boldsymbol{x}(0)$ is chosen s.t. $\mathbf{1}^{\mathrm{T}}\boldsymbol{x}(0) = 1$, it holds:

$$\boldsymbol{x}(n+1) = (1-\tau)\mathbf{A}\boldsymbol{x}(n) + \frac{\tau}{|\mathcal{V}|}\mathbf{1}.$$

Above dynamic is of the form (9.6) with $\mathbf{B} = (1-\tau)\mathbf{A}$ and $\boldsymbol{u} = (\tau/|\mathcal{V}|)\mathbf{1}$.

Example 19 (Randomized opinion model with informational bias): At time $k = 0$, each agent $i \in [N]$ starts with an initial opinion $\theta_i(0)$. For time k:

$$\theta_i(k) = \begin{cases} \sum_j a_{ij}(k)Y_j(k), & \text{w.p. } \gamma_i \\ \theta_i(0), & \text{w.p. } 1-\gamma_i \end{cases},$$

where $Y_j(k)$, j, k, independent RVs with $\mathbb{E}[Y_j(k)|\theta_j(k-1)] = \theta_j(k-1)$. It holds:

$$\mathbb{E}[\theta_i(k)] = \gamma_i \sum_j a_{i,j}(k)\mathbb{E}[\theta_j(k-1)] + (1-\gamma_i)\mathbb{E}[\theta_i(0)].$$

So the dynamic of the expectation of the agents' opinion coincide with the Johnsen-Friedkin model (9.3)

9.2.6. Cluster Structure

In this work, we are interested in the opinion dynamic of clustered agents. Specifically, we consider the case where the population of agents can be segregated into $M \in \mathbb{N}$ clusters $[M]$, where each cluster i contains $K_i \in \mathbb{N}$ agents with:

$$\sum_{i=1}^{M} K_i = N.$$

By eventually reindexing the agents, we may write the given trust matrix as:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1M} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{M1} & \cdots & \mathbf{A}_{MM} \end{bmatrix},$$

where $\mathbf{A}_{ii} \in \mathbb{R}^{K_i \times K_i}$ denotes the intra-cluster communication matrix within the cluster i and $\mathbf{A}_{ij} \in \mathbb{R}^{K_i \times K_j}$ denotes the matrix which specifies how cluster j influenced cluster i. In the case where the inter-cluster communications are negligible, i.e., where the entries of $\mathbf{A}_{ij}$ is small, we speak of clustered structure. Now given a cluster $i \in [M]$. We denote the set of agents not in cluster i by $[\backslash i]$.

9.3. Uncoupling the opinion's Dynamic

The aim of this section is to introduce the notion of substochastic complementation which allows one to effectively uncouple the Friedkin-Johnsen dynamic in subdynamics each run independently in the corresponding clusters. In order to make our approach transparent we first recall the notion of stochastic complementation developed in the seminal work [201]

9.3.1. Stochastic Complementation

In case that the population of agents possess a clustered structure, We define the stochastic complementation of $\mathbf{A}$ by:

$$\mathbf{C} := \begin{bmatrix} \mathbf{C}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{C}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{C}_M \end{bmatrix},$$

where:

$$\mathbf{C}_i = \mathbf{A}_{ii} + \mathbf{A}_{i:}(\mathbf{I} - \mathbf{A}_{\backslash i})^{-1}\mathbf{A}_{:i},$$

with:

$$\mathbf{A}_{i:} := [\mathbf{A}_{i,1} \cdots \mathbf{A}_{i,(i-1)} \ \mathbf{A}_{i,(i+1)} \cdots \mathbf{A}_{i,M}] \in \mathbb{R}^{K_i \times N-K_i},$$

and:

$$\mathbf{A}_{:i}^{\mathrm{T}} := [\mathbf{A}_{1,i}^{\mathrm{T}} \cdots \mathbf{A}_{(i-1),i}^{\mathrm{T}} \ \mathbf{A}_{(i+1),i}^{\mathrm{T}} \cdots \mathbf{A}_{M,i}^{\mathrm{T}}] \in \mathbb{R}^{K_i \times (N-K_i)},$$

and $\mathbf{A}_{\backslash i} \in \mathbb{R}^{(N-K_i)\times(N-K_i)}$ is the matrix obtained by deleting the i-th row and column of blocks. So above matrix describes the case where the opinion of an agent in clusters i exit the latter, subsequently staying (for possibly infinite times), and going back to cluster i.

Several properties of of $\mathbf{A}$ transfer to its stochastic complementation (see [201]):

Lemma 9.8: *Suppose that* $\mathbf{A}$ *is an irreducible stochastic matrix. Then* $\mathbf{C}$ *is well-defined. Moreover, for any* i, $\mathbf{C}_i$ *is an irreducible stochastic matrix*

It holds:

$$\mathbf{A}_{i:}(\mathbf{I} - \mathbf{A}_{\backslash i})^{-1}\mathbf{A}_{:i} = \sum_{k=0}^{\infty} \mathbf{A}_{i:}\mathbf{A}_{\backslash i}^{k}\mathbf{A}_{i:}.$$

The following statement is known:

Theorem 9.9: *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be an irreducible stochastic matrix partitioned into* M^2

block matrices each of size $K \times K$. It holds:

$$\|\mathbf{A} - \mathbf{C}\|_\infty \leqslant 2 \max_i \|\mathbf{A}_{i:}\|_\infty$$

Since the original proof in [201] contains a slight error, we provide the proof of above statement in the following:

Proof (Proof of Theorem 9.9): We compute:

$$\begin{aligned}
\|\mathbf{A} - \mathbf{C}\|_\infty &= \max_i \| \left[\mathbf{A}_{i,1}, \mathbf{A}_{i,2}, \dots, \mathbf{A}_{ii} - \mathbf{C}_i, \dots, \mathbf{A}_{i,M}\right] \|_\infty \\
&= \max_i \|\mathbf{A}_{i,1}\mathbf{1}_{K_1} + \dots + |\mathbf{A}_{ii} - \mathbf{C}_i|\,\mathbf{1}_{K_i} + \dots + \mathbf{A}_{i,M}\mathbf{1}_{K_M}\|_\infty \\
&= \max_i \| \sum_{j \neq i} \mathbf{A}_{ij}\mathbf{1}_{K_j} + |\mathbf{A}_{ii} - \mathbf{C}_i|\,\mathbf{1}_{K_i}\|_\infty \\
&= \max_i \|\mathbf{A}_{i:}\mathbf{1}_{N-K_i} + |\mathbf{A}_{ii} - \mathbf{C}_i|\,\mathbf{1}_{K_i}\|_\infty,
\end{aligned}$$

where for a matrix $\mathbf{B}$, $|\mathbf{B}|$ denotes the matrix whose entries are the absolute value of the entries of $\mathbf{B}$. Now we aim to compute:

$$|\mathbf{A}_{ii} - \mathbf{C}_i|\,\mathbf{1}_{K_i} = \mathbf{A}_{i:}(\mathbf{I}_{N-K_i} - \mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i}\mathbf{1}_{K_i}, \tag{9.13}$$

where the equality follows from the fact that the entries of $\mathbf{A}_{i:}(\mathbf{I}_{N-K} - \mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i}$ are non-negative. To this end, observe that since $\mathbf{A}$ is stochastic:

$$\mathbf{A}_{\setminus i}\mathbf{1}_{N-K} + \mathbf{A}_{:i}\mathbf{1}_K = \mathbf{1}_{N-K}.$$

Consequently:

$$\mathbf{A}_{:i}\mathbf{1}_{K_i} = (\mathbf{I}_{N-K_i} - \mathbf{A}_{\setminus i})\mathbf{1}_{N-K_i},$$

and thus:

$$(\mathbf{I}_{N-K_i} - \mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i}\mathbf{1}_{K_i} = \mathbf{1}_{N-K_i}.$$

Setting this into (9.13):

$$(\mathbf{A}_{ii} - \mathbf{C}_i)\mathbf{1}_{K_i} = -\mathbf{A}_{i:}\mathbf{1}_{N-K_i}.$$

Finally, we have as desired:

$$\|\mathbf{A} - \mathbf{C}\|_\infty = \max_i \|2\mathbf{A}_{i:}\mathbf{1}_{N-K}\|_\infty = 2 \max_i \|\mathbf{A}_{i:}\|_\infty \qquad \blacksquare$$

Above bound leads to the following result:

Theorem 9.10: *We have:*

$$\|\mathbf{A}^n - \mathbf{C}^n\|_\infty \leqslant n\delta, \quad \text{where } \delta := 2 \max_i \|\mathbf{A}_{i:}\|_\infty,$$

and consequently:

$$\| \mathrm{St}[\mathbf{A}, 0, \boldsymbol{x}_0](k) - \mathrm{St}[\mathbf{C}, 0, \boldsymbol{x}_0](k)\|_\infty \leqslant n\delta\|\boldsymbol{x}_0\|_\infty$$

Above Theorem gives a guarantee that for small times, the undisturbed opinion dynamic, behaves as if the clusters were isolated, since it is similar to the benchmark dynamic $\mathrm{St}[\mathbf{C}, 0, x_0]$ containing cluster-individual dynamics. Moreover, above Theorem supports the intuition that in small times the opinion of each individual propagates first within its cluster before reaching the whole network. However, one cannot expect that the clusterness of the population's opinion is true for large times, since otherwise the opinion mixing property in asymptotic regime (Proposition 9.1) would be violated.

9.3.2. Substochastic Complementation

Now in case that $\mathbf{\Gamma} \neq \mathbf{I}$, we not only have the communication matrix $\mathbf{A}$ as a factor determining the opinion dynamic, but rather a mixture of the communication matrix $\mathbf{A}$ and disturbance sequences $\xi(k)$. So analyzing the corresponding opinion dynamic by means of the stochastic complementation seems to be inappropriate since it only considers the underlying structure of the network specified by $\mathbf{A}$, while the extrinsic influence/bias, specified by $\mathbf{\Gamma}$ remains ignored.

In order to overcome this drawback, we introduce the notion of substochastic complementation. For this sake, let us first define:

$$\mathbf{\Gamma}_{\setminus i} = \mathrm{diag}(\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_{i-1}, \mathbf{\Gamma}_{i+1}, \dots, \mathbf{\Gamma}_M) \in \mathbb{R}^{(M-K_i)\times(M-K_i))},$$

We define the substochastic complementation of the tuple $(\mathbf{A}, \mathbf{\Gamma})$ (or shorter: $\mathbf{\Gamma A}$) w.r.t. cluster i as:

$$\tilde{\mathbf{C}}_i = \mathbf{\Gamma}_i \mathbf{A}_{ii} + \mathbf{\Gamma}_i \mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}.$$

In order that $\tilde{\mathbf{C}}_i$ is well defined, we need to ensure that the Neumann series $(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}$ is well-defined. The latter holds if and only if $\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}$ is Schur-stable. It is well-known that the Schur-stability of $\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}$ is ensured if $\mathbf{\Gamma A}$ is an irreducible proper substochastic matrix. This is a consequence of the following Lemma (see e.g. [203]):

Lemma 9.11: *Let $\mathbf{B}$ be a non-negative matrix. For any principle submatrix $\tilde{\mathbf{B}}$ of $\mathbf{B}$, we have $\rho(\tilde{\mathbf{B}}) \leqslant \rho(\tilde{\mathbf{B}})$. If in addition $\mathbf{B}$ is irreducible and $\tilde{\mathbf{B}} \neq \mathbf{B}$, then $\rho(\tilde{\mathbf{B}}) < \rho(\tilde{\mathbf{B}})$*

We can emphasize the cluster structure induced by the substochastic complementation by defining the block diagonal matrix:

$$\tilde{\mathbf{C}} := \begin{bmatrix} \tilde{\mathbf{C}}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\mathbf{C}}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mathbf{C}}_M \end{bmatrix},$$

Several properties of $\mathbf{\Gamma A}$ transfer to its substochastic complementation:

Proposition 9.12: *We have:*

1. $\tilde{\mathbf{C}}_i$ *is substochastic*

2. *If* $\mathbf{\Gamma A}$ *is irreducible, then* $\tilde{\mathbf{C}}_i$ *is irreducible*

3. *If there exists* $j \in \mathcal{I}_i$ *s.t.* $\gamma_j = 0$*, then* $\tilde{\mathbf{C}}_i$ *is reducible*

In order to show above properties, we need the following preliminary result:

Lemma 9.13: *It holds:*

$$[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\mathbf{1} = \mathbf{1} + [\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{\Gamma}_{\setminus i} - \mathbf{I}]\mathbf{1}$$

Proof: We have since $\mathbf{A}$ is stochastic:

$$\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\mathbf{1} + \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}\mathbf{1} = \mathbf{\Gamma}_{\setminus i}\mathbf{1} = \mathbf{1} + (\mathbf{\Gamma}_{\setminus i} - \mathbf{I})\mathbf{1},$$

and thus:

$$\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\mathbf{1} = -\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}\mathbf{1} + \mathbf{1} + (\mathbf{\Gamma}_{\setminus i} - \mathbf{I})\mathbf{1} = (\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})\mathbf{1} + (\mathbf{\Gamma}_{\setminus i} - \mathbf{I})\mathbf{1}.$$

Multiplying both sides of the above identity by $(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}$, we obtain the desired statement. ■

Now, we are ready to provide the proof of Proposition 9.12:

Proof (Proof of Proposition 9.12): We first show that $\tilde{\mathbf{C}}_i$ is substochastic. We have:

$$\tilde{\mathbf{C}}_i\mathbf{1} = \mathbf{\Gamma}_i(\mathbf{A}_{ii} + \mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i})\mathbf{1}.$$

Clearly $\mathbf{A}_{ii}\mathbf{1} \leqslant \mathbf{1}$. Moreover by Lemma 9.13, it follows:

$$\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\mathbf{1} = \mathbf{A}_{i:}(\mathbf{I} - (\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I} - \mathbf{\Gamma}_i))\mathbf{1} \leqslant \mathbf{A}_{:i}\mathbf{1},$$

where the inequality follows from the fact that $(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I} - \mathbf{\Gamma}_i)$ is non-negative.

Setting both estimates into (58), we obtain:

$$\tilde{\mathbf{C}}_i \mathbf{1} \leqslant \mathbf{\Gamma}_i(\mathbf{A}_{ii}\mathbf{1} + \mathbf{A}_{i:}\mathbf{1}) = \mathbf{\Gamma}_i, \tag{9.14}$$

where the last equality follows from stochasticity of $\mathbf{A}$. Since $\mathbf{\Gamma}_i \mathbf{1} \leqslant \mathbf{1}$, we obtain the desired fact that $\tilde{\mathbf{C}}_i$ is substochastic.

Now we want to show that $\mathbf{\Gamma A}$ is irreducible implies that $\tilde{\mathbf{C}}_i$ is irreducible. Toward this end, let $j, k \in \mathcal{I}_i$, where $\mathcal{I}_i$ is the index set of agents in cluster i. We want to show that $[\tilde{\mathbf{C}}_i^m]_{jk} > 0$ for an $m \in \mathbb{N}$. Since $\mathbf{\Gamma A}$ is irreducible, we know that there exists a path from j to k in $\mathcal{G}(\mathbf{\Gamma A})$. First, suppose that there is a path from j to k in $\mathcal{G}(\mathbf{\Gamma}_i \mathbf{A}_{ii})$, then it follows that:

$$[\mathbf{\Gamma}_i \mathbf{A}_{ii}]_{jk}^m > 0,$$

for an $m \in \mathbb{N}$. Consequently:

$$[\tilde{\mathbf{C}}_i^m]_{jk} \geqslant [(\mathbf{\Gamma}_i \mathbf{A}_{ii})^m]_{jk} > 0.$$

Now suppose that there is no path from j to k in $\mathcal{G}(\mathbf{\Gamma}_i \mathbf{A}_{ii})$. Irreducibility of $\mathbf{A}$ asserts that there must be a path going through $\mathcal{I}_i^c$. Specifically, we have a path:

$$\begin{aligned} j &\to j_1^{(1)} \to \cdots \to j_{p_1}^{(1)} \to q_1^{(1)} \to \cdots \to q_{r_1}^{(1)} \to j_1^{(2)} \to \cdots \to j_{p_2}^{(2)} \to q_1^{(2)} \to \cdots \to q_{r_2}^{(2)} \\ &\to \cdots \to j_1^{(s)} \to \cdots \to j_{p_s}^{(s)} \to k, \end{aligned}$$

where $s \in \mathbb{N}$, $(j_b^{(1)})_{b\in[p_1]}, \ldots, (j_b^{(s)})_{b\in[p_s]}$ are in $\mathcal{I}_i$, and $(q_b^{(1)})_{b\in[r_1]}, \ldots, (q_b^{(s-1)})_{b\in[r_{s-1}]}$ are in $\mathcal{I}_i^c$. Denote $\mathbf{B} = \mathbf{\Gamma A}$. This asserts that:

$$[\mathbf{B}_{ii}^{p_1-1}\mathbf{B}_{i:}\mathbf{B}_{\backslash i}^{r_1-1}\mathbf{B}_{:i}\mathbf{B}_{ii}^{p_2-1}\mathbf{B}_{i:}\mathbf{B}_{\backslash i}^{r_2-1}\mathbf{B}_{:i}\cdots\mathbf{B}_{ii}^{p_s-1}]_{jk} > 0.$$

Noticing that $(\mathbf{I} - \mathbf{B}_{\backslash i})^{-1} = \sum_{h=0}^{\infty} \mathbf{B}_{\backslash i}^h$, we have that:

$$\mathbf{B}_{ii}^{p_1-1}\mathbf{B}_{i:}\mathbf{B}_{\backslash i}^{r_1-1}\mathbf{B}_{:i}\mathbf{B}_{ii}^{p_2-1}\mathbf{B}_{i:}\mathbf{B}_{\backslash i}^{r_2-1}\mathbf{B}_{:i}\cdots\mathbf{B}_{ii}^{p_s-1}$$

is a summand in $\tilde{\mathbf{C}}_i^m$ for a certain $m \in \mathbb{N}$. Consequently, we have as desired:

$$[\tilde{\mathbf{C}}_i^m]_{jk} \geqslant [\mathbf{B}_{ii}^{p_1-1}\mathbf{B}_{i:}\mathbf{B}_{\backslash i}^{r_1-1}\mathbf{B}_{:i}\mathbf{B}_{ii}^{p_2-1}\mathbf{B}_{i:}\mathbf{B}_{\backslash i}^{r_2-1}\mathbf{B}_{:i}\cdots\mathbf{B}_{ii}^{p_s-1}]_{jk} > 0.$$

In order to show the last statement, notice that $\tilde{C}_i = \mathbf{\Gamma}_i \tilde{\mathbf{B}}$, for a matrix $\tilde{\mathbf{B}}$. Therefore, if there exists $j \in \mathcal{I}_i$ s.t. $\gamma_j = 0$, then it follows that the j-th row of $\tilde{\mathbf{C}}_i$ is zero yielding the desired statement. ∎

We can provide an interpretation of the substochastic complement by means of the following notion:

Definition 9.1 (Schur Complement): *Suppose that* $\mathbf{B}_{11}$, $\mathbf{B}_{12}$, $\mathbf{B}_{21}$, $\mathbf{B}_{22}$ *are* $p \times p$,$p \times$

q, $q \times p$, and $q \times q$ matrices, and that $\mathbf{B}_{22}$ is invertible. Define:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}.$$

Then the Schur complement of the block $\mathbf{B}_{22}$ in the matrix $\mathbf{B}$ is defined as the matrix:

$$\mathbf{B}/\mathbf{B}_{22} := \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21}$$

In order to relate the notion of the substochastic complementation with the notion of Schur complement, we first define $\mathbf{Q}_{i\leftrightarrow 1}$ as the permutation matrix interchanging from left the i-th and the 1st row of a matrix (resp. from right the column). It holds:

$$\mathbf{Q}_{i\leftrightarrow 1}\mathbf{\Gamma}\mathbf{A}\mathbf{Q}_{i\leftrightarrow 1} = \begin{bmatrix} \mathbf{\Gamma}_i\mathbf{A}_{ii} & \mathbf{\Gamma}_i\mathbf{A}_{i:} \\ \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i} & \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i} \end{bmatrix}.$$

Now, since:

$$\mathbf{I} - \mathbf{Q}_{i\leftrightarrow 1}\mathbf{\Gamma}\mathbf{A}\mathbf{Q}_{i\leftrightarrow 1} = \begin{bmatrix} \mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii} & -\mathbf{A}_{i:} \\ -\mathbf{A}_{:i} & \mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i} \end{bmatrix},$$

we have:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i) = \mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii}-\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i} = (\mathbf{I}-\mathbf{Q}_{i\leftrightarrow 1}\mathbf{\Gamma}\mathbf{A}\mathbf{Q}_{i\leftrightarrow 1})/(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}). \quad (9.15)$$

Therefore we have (up to a permutation):

$(\mathbf{I}-$[Substochastic Complement of $(\mathbf{\Gamma}, \mathbf{A})$ $)]_{ii}$ = Schur Complement of $\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}$ in $\mathbf{I} - \mathbf{\Gamma}\mathbf{A}$.

This observation will be important for our later approaches.

9.4. Bounds for the Discrepancy of the Matrix Approximation

In this section, we derive measures which quantify the error of the approximation of the Friedkin-Johnsen dynamic by the dynamic induced by the substochastic complementation and thus also the clusterness of the opinion dynamic. Throughout the remaining of this work, we assume the following:

Assumption 9.1: $\mathbf{\Gamma}\mathbf{A}$ *is Schur-stable.*

As we already discuss, this implies that the substochastic complementation exists and also that the Friedkin-Johnsen dynamic has a limit. The latter is important for our investigations of the behaviour of the Friedkin-Johnsen dynamic in large times done in the end of this section.

9.4.1. One Shot Bound

At first we aim to derive the following bound for the error of the approximation of the weighted communication matrix by means of its stochastic complementation. Toward this direction, notice first that by the block structure of $\tilde{\mathbf{C}}$ and the definition of the infinite norm of a matrix, we have:

$$\|\tilde{\mathbf{C}} - \mathbf{\Gamma}\mathbf{A}\|_\infty = \max_{i \in [M]} \| \left| \tilde{\mathbf{C}}_i - \mathbf{\Gamma}_i \mathbf{A}_{ii} \right| \mathbf{1} + \mathbf{\Gamma}_i \mathbf{A}_{i:} \mathbf{1} \|_\infty .$$

Let now $i \in [M]$ be arbitrary. We have:

$$\left| \tilde{\mathbf{C}}_i - \mathbf{\Gamma}_i \mathbf{A}_{ii} \right| = \mathbf{\Gamma}_i \mathbf{A}_{i:} (\mathbf{I} - \mathbf{\Gamma}_{\backslash i} \mathbf{A}_{\backslash i})^{-1} \mathbf{\Gamma}_{\backslash i} \mathbf{A}_{:i},$$

and consequently:

$$\|\tilde{\mathbf{C}} - \mathbf{\Gamma}\mathbf{A}\|_\infty = \max_{i \in [M]} \|\mathbf{\Gamma}_i \mathbf{A}_{i:} \mathbf{1} + \mathbf{\Gamma}_i \mathbf{A}_{i:} (\mathbf{I} - \mathbf{\Gamma}_{\backslash i} \mathbf{A}_{\backslash i})^{-1} \mathbf{\Gamma}_{\backslash i} \mathbf{A}_{:i} \mathbf{1}\|_\infty . \tag{9.16}$$

From this identity, we can induce a bound for $\|\tilde{\mathbf{C}}^n - \mathbf{\Gamma}\mathbf{A}^n\|$. The main ingredient is the following lemma:

Lemma 9.14: *Given two sequences of $D \times D$ real matrices $(\mathbf{B}(n))_{n \in \mathbb{N}}$ and $(\tilde{\mathbf{B}}(n))_{n \in \mathbb{N}}$. For any $k \in \mathbb{N}$, it holds:*

$$\mathbf{B}(1:k) - \tilde{\mathbf{B}}(1:k) = \sum_{i=1}^{k} \tilde{\mathbf{B}}(i+1:k)[\mathbf{B}(i) - \tilde{\mathbf{B}}(i)]\mathbf{B}(1:i-1), \tag{9.17}$$

where for a sequence of $D \times D$ matrices $(\mathbf{A}(n))_{n \in \mathbb{N}}$, we denote:

$$\mathbf{A}(k:l) = \mathbf{A}(k)\mathbf{A}(k+1)\cdots\mathbf{A}(l), \quad k \leqslant l$$

Proof: Note that for $k = 1$, the statement is true. So suppose that it is true for a $k \in \mathbb{N}$. We have induction assumption:

$$\begin{aligned}
&\sum_{i=1}^{k+1} \tilde{\mathbf{B}}(i+1:k+1)[\mathbf{B}(i) - \tilde{\mathbf{B}}(i)]\mathbf{B}(1:i-1) \\
&= [\mathbf{B}(k+1) - \tilde{\mathbf{B}}(k+1)]\mathbf{B}(1:k) + \tilde{\mathbf{B}}(k+1) \sum_{i=1}^{k} \tilde{\mathbf{B}}(i+1:k)[\mathbf{B}(i) - \tilde{\mathbf{B}}(i)]\mathbf{B}(1:i-1) \\
&= [\mathbf{B}(k+1) - \tilde{\mathbf{B}}(k+1)]\mathbf{B}(1:k) + \tilde{\mathbf{B}}(k+1)[\mathbf{B}(1:k) - \tilde{\mathbf{B}}(1:k)] \\
&= \mathbf{B}(1:k+1) - \tilde{\mathbf{B}}(k+1)\mathbf{B}(1:k) + \tilde{\mathbf{B}}(k+1)\mathbf{B}(1:k) - \tilde{\mathbf{B}}(1:k+1) \\
&= \mathbf{B}(1:k+1) - \tilde{\mathbf{B}}(1:k+1),
\end{aligned}$$

as desired. ∎

As a consequence of above Lemma, we have:

$$\begin{aligned}
\|\tilde{\mathbf{C}}^n - (\mathbf{\Gamma A})^n\|_\infty &= \|\sum_{k=1}^{n} \tilde{\mathbf{C}}^{n-k+1}(\tilde{\mathbf{C}} - \mathbf{\Gamma A})(\mathbf{\Gamma A})^{k-1}\|_\infty \\
&\leqslant \sum_{k=1}^{n} \|\tilde{\mathbf{C}}^{n-k+1}\|_\infty \|\tilde{\mathbf{C}} - \mathbf{\Gamma A}\|_\infty \|(\mathbf{\Gamma A})^{k-1}\|_\infty \qquad (9.18) \\
&\leqslant \sum_{k=1}^{n} \|\tilde{\mathbf{C}} - \mathbf{\Gamma A}\|_\infty = n\|\tilde{\mathbf{C}} - \mathbf{\Gamma A}\|_\infty .
\end{aligned}$$

The first inequality in above computation follows from the triangle inequality, the second from the fact that $\mathbf{\Gamma A}$ and $\tilde{\mathbf{C}}$ are substochastic.

Having now an estimate for $\|\tilde{\mathbf{C}}^n - (\mathbf{\Gamma A})^n\|_\infty$ by (9.16), we can obtain the same for the error of the cluster approximation of the Friedkin-Johnsen dynamic:

$$\begin{aligned}
&\| \mathrm{St}[\mathbf{\Gamma A}, \mathbf{\Lambda}\xi, x_0](n) - \mathrm{St}[\tilde{\mathbf{C}}, \mathbf{\Lambda}\xi, x_0](n)\|_\infty = \|[(\mathbf{\Gamma A})^n - \tilde{\mathbf{C}}^n]\boldsymbol{x}_0 + [\sum_{k=0}^{n-1}(\mathbf{\Gamma A})^k\mathbf{\Lambda} - \sum_{k=0}^{n-1}\tilde{\mathbf{C}}^k\mathbf{\Lambda}]\xi\|_\infty \\
&\leqslant \|(\mathbf{\Gamma A})^n - \tilde{\mathbf{C}}^n\|_\infty\|\boldsymbol{x}_0\|_\infty + \|[\sum_{k=0}^{n-1}(\mathbf{\Gamma A})^k\mathbf{\Lambda} - \sum_{k=0}^{n-1}\tilde{\mathbf{C}}^k\mathbf{\Lambda}]\|_\infty\|\xi\|_\infty \\
&\leqslant \|(\mathbf{\Gamma A})^n - \tilde{\mathbf{C}}^n\|_\infty\|\boldsymbol{x}_0\|_\infty + \sum_{k=0}^{n-1}\|(\mathbf{\Gamma A})^k - \tilde{\mathbf{C}}^k\|_\infty\|\mathbf{\Lambda}\|_\infty\|\xi\|_\infty \leqslant n\delta + \frac{(n-1)n}{2}\delta\|\mathbf{\Lambda}\|_\infty\|\xi\|_\infty,
\end{aligned}$$
(9.19)

where:

$$\delta := \max_{i\in[M]} \|\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\mathbf{1}\|_\infty$$

Remark 23: Notice that above bound has a strong dependency on the time n (of order $\mathcal{O}(n^2)$). However, by further computations, one can weaken this dependency. For instance, one can tighten the estimate (9.18), by providing an upper bound for the power of matrices $\|\tilde{\mathbf{C}}^{n-k+1}\|_\infty$ $\|(\mathbf{\Gamma A})^{k-1}\|_\infty$ better than one. For sufficiently large power, this is possible since the matrices are (proper) substochastic.

At last, we want to show that $\tilde{\mathbf{C}}$ gives a better approximation for $\mathbf{\Gamma A}$ than that induced by stochastic complementation, i.e. $\mathbf{\Gamma C}$. meaning that $\|\tilde{\mathbf{C}} - \mathbf{\Gamma A}\|_\infty$ is smaller than the bound $\max_i \|2\mathbf{\Gamma}_i\mathbf{A}_{i:}\|_\infty$ for $\|\mathbf{\Gamma C} - \mathbf{\Gamma A}\|_\infty$ obtained by analogous technique as used for showing the Theorem 9.9. A consequence of Lemma 9.13 is:

$$\begin{aligned}
\left|\tilde{\mathbf{C}}_i - \mathbf{\Gamma}_i\mathbf{A}_{ii}\right|\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} &= \mathbf{\Gamma}_i\mathbf{A}_{i:}\left[\mathbf{1} + [\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{\Gamma}_{\setminus i} - \mathbf{I}]\right]\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} \\
&= 2\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} - \mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}]\mathbf{1},
\end{aligned}$$

so that:

$$\begin{aligned}\|\tilde{\mathbf{C}} - \mathbf{\Gamma A}\|_\infty &= \max_i \|2\mathbf{\Gamma}_i\mathbf{A}_{i:} - \mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}]\|_\infty \\ &= \max_{i\in[N]} \max_j \left[\mathbf{\Gamma}_i\mathbf{A}_{i:}\left(2\mathbf{I} - [\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}]\right)\mathbf{1}\right]_j .\end{aligned} \tag{9.20}$$

Now, since $\mathbf{\Gamma}_i\mathbf{A}_{i:} \geqslant 0$, we have that:

$$\mathbf{\Gamma}_i\mathbf{A}_{i:}\left(2\mathbf{I} - [\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}]\right)\mathbf{1} \leqslant 2\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1}.$$

Implying that the given bound for $\|\tilde{\mathbf{C}} - \mathbf{\Gamma A}\|_\infty$ is smaller than the bound $\|2\mathbf{\Gamma}_i\mathbf{A}_{i:}\|_\infty$ for $\|\mathbf{\Gamma C} - \mathbf{\Gamma A}\|_\infty$ giving a hint that $\tilde{\mathbf{C}}$ is a better benchmark than $\mathbf{\Gamma C}$. Moreover, the fact that $[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}]\mathbf{1} \leqslant 1$ implies that

$$\mathbf{\Gamma}_i\mathbf{A}_{i:}\left(2\mathbf{I} - [\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I} - \mathbf{\Gamma}_{\setminus i}]\right)\mathbf{1} \geqslant \mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1}.$$

So the improvement we can obtain is the factor 1 which has a non-negligible impact in the non-asymptotic analysis as done in (9.19).

9.4.2. Bound for infinite accumulation

The drawback of the analysis based on the one-shot bound (9.16) is that the corresponding estimate may depends on the time. So it is not suited for forecasting the system behaviour in large times. Motivated by the goal to eliminate this drawback, our aim in this section is to analyze the quantity:

$$\mathbf{h} := \left|(\mathbf{I} - \mathbf{\Gamma A})^{-1}(\mathbf{I} - \mathbf{\Gamma}) - (\mathbf{I} - \tilde{\mathbf{C}})^{-1}(\mathbf{I} - \mathbf{\Gamma})\right|\mathbf{1}, \tag{9.21}$$

Before we continue with our discussion, we need first to show that $\tilde{\mathbf{C}}_i$ is Schur-stable:

Lemma 9.15: *If $\mathbf{\Gamma A}$ is Schur-stable, then $\tilde{\mathbf{C}}_i$ is also Schur-stable.*

Proof: From (9.15) and the well-known formula specifying the determinant of a Schur complement:

$$\det(\mathbf{B}) = \det(\mathbf{B}/\tilde{\mathbf{B}})\det(\tilde{\mathbf{B}}),$$

we have:

$$\det(\mathbf{I}-\mathbf{\Gamma A}) = \det(\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\mathbf{\Gamma A})\mathbf{Q}_{i\leftrightarrow 1}) = \det(\mathbf{I}-\mathbf{Q}_{i\leftrightarrow 1}\mathbf{\Gamma A}\mathbf{Q}_{i\leftrightarrow 1}) = \det(\tilde{\mathbf{C}}_i)\det(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}), \tag{9.22}$$

where the first equality follows from the fact that the determinant is invariant of simultaneous row and column permutations. Now, we have that $\det(\mathbf{I} - \mathbf{\Gamma A}) \neq 0$ since $\mathbf{\Gamma A}$ is Schur-stable. Moreover, by the inequality $\rho(\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}) \leqslant \rho(\mathbf{\Gamma A})$ we have that $\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}$ is Schur-stable and consequently $\det(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}) \neq 0$. From (9.22), we can now infer that

$(\mathbf{I}-\tilde{\mathbf{C}}_i)$ and thus the desired statement. ■

Now, in order to see the meaning of $\mathbf{h}$, notice first that by (9.7), it holds:

$$\left|\mathrm{St}[\mathbf{\Gamma A}, \mathbf{\Lambda}\xi, \boldsymbol{x}_0](n) - \mathrm{St}[\tilde{\mathbf{C}}, \mathbf{\Lambda}\xi, \boldsymbol{x}_0](n)\right| = \left|[(\mathbf{\Gamma A})^n - \tilde{\mathbf{C}}^n]\boldsymbol{x}_0 + [\sum_{k=0}^{n-1}(\mathbf{\Gamma A})^k\mathbf{\Lambda} - \sum_{k=0}^{n-1}\tilde{\mathbf{C}}^k\mathbf{\Lambda}]\xi\right|$$

Both $\mathbf{\Gamma A}$ and $\tilde{\mathbf{C}}$ are Schur-stable, giving:

$$(\mathbf{\Gamma A})^n, \tilde{\mathbf{C}}^n \xrightarrow{n\to\infty}, \ \sum_{k=0}^{n-1}(\mathbf{\Gamma A})^k \xrightarrow{n\to\infty} (\mathbf{I}-\mathbf{\Gamma A})^{-1}, \text{ and } \sum_{k=0}^{n-1}\tilde{\mathbf{C}}^k \xrightarrow{n\to\infty} (\mathbf{I}-\tilde{\mathbf{C}})^{-1},$$

and consequently:

$$\lim_{n\to\infty}\left|\mathrm{St}[\mathbf{\Gamma A}, \mathbf{\Lambda}\xi, \boldsymbol{x}_0](n) - \mathrm{St}[\tilde{\mathbf{C}}, \mathbf{\Lambda}\xi, \boldsymbol{x}_0](n)\right| = \left|\left[(\mathbf{I}-\mathbf{\Gamma A})^{-1}(\mathbf{I}-\mathbf{\Gamma}) - (\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})\right]\xi\right|.$$

From above identity, it follows that $\mathbf{h}$ gives rise to the worst-case asymptotic error of the approximation of the Friedkin-Johnsen dynamic (9.4) by means of the cluster-localized dynamic $\mathrm{St}[\tilde{\mathbf{C}}, \mathbf{\Gamma}\xi, \boldsymbol{x}_0]$, i.e., we have:

$$\lim_{n\to\infty}\left|\mathrm{St}[\mathbf{\Gamma A}, \mathbf{\Lambda}\xi, \boldsymbol{x}_0](n) - \mathrm{St}[\tilde{\mathbf{C}}, \mathbf{\Lambda}\xi, \boldsymbol{x}_0](n)\right| \leqslant \mathbf{h}\|\xi\|_\infty.$$

Our main result is the following:

Theorem 9.16: $\mathbf{h}$ *has non-negative entries and can be uncoupled as follows*

$$\mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_M \end{bmatrix},$$

where:

$$\begin{aligned} \mathbf{h}_i &= (\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1} && (9.23)\\ &= (\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{\Gamma}_i-\tilde{\mathbf{C}}_i)\mathbf{1} && (9.24)\\ &= \mathbf{1}-(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1} && (9.25)\end{aligned}$$

In order to show above result, we first need some auxiliary statements:

Lemma 9.17: *The i-th block diagonal element of* $(\mathbf{I}-\mathbf{\Gamma A})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ *is* $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)$ *and the off block-diagonal part of the i-th block-row is a block-column permutation of* $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})$

Proof: For ease of notation, we define $\mathbf{B} := \mathbf{\Gamma}\mathbf{A}$. Analogous to $\mathbf{A}$, we part $\mathbf{B}$ as:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{1,1} & \cdots & \mathbf{B}_{1,M} \\ \vdots & \ddots & \vdots \\ \mathbf{B}_{M,1} & \cdots & \mathbf{B}_{M,M} \end{bmatrix},$$

where $\mathbf{B}_{ij} = \mathbf{\Gamma}_i\mathbf{A}_{ij}$. Now we take a permutation matrix $\mathbf{Q}_{i\leftrightarrow 1}$ corresponding to the interchange of the first and the i-th block rows (or also column). We have:

$$\hat{\mathbf{B}}^{(i)} := \mathbf{Q}_{i\leftrightarrow 1}\mathbf{B}\mathbf{Q}_{i\leftrightarrow 1} = \begin{bmatrix} \mathbf{B}_{i,i} & \mathbf{B}_{i,:} \\ \mathbf{B}_{:,i} & \mathbf{B}_{\backslash i} \end{bmatrix}. \tag{9.26}$$

Moreover by setting $\mathbf{B} = \mathbf{I}$ in above identity, we have also:

$$\mathbf{Q}_{i\leftrightarrow 1}\mathbf{Q}_{i\leftrightarrow 1} = \mathbf{I}. \tag{9.27}$$

Now, (9.26) asserts:

$$\mathbf{I} - \hat{\mathbf{B}}^{(i)} = \begin{bmatrix} \mathbf{I} - \hat{\mathbf{B}}_{i,i} & -\hat{\mathbf{B}}_{i,:} \\ -\hat{\mathbf{B}}_{:,i} & \mathbf{I} - \hat{\mathbf{B}}_{\backslash i} \end{bmatrix}.$$

Applying the formula of matrix inversion in block form and subsequently the identity:

$$(\mathbf{I}-\mathbf{B}_{i,i}) - (-\hat{\mathbf{B}}_{i,:})(\mathbf{I}-\hat{\mathbf{B}}_{\backslash i})^{-1}(-\hat{\mathbf{B}}_{:,i}) = \mathbf{I} - (\mathbf{\Gamma}_i\mathbf{A}_{ii} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{:i}) = \mathbf{I} - \tilde{\mathbf{C}}_i,$$

we have:

$$\begin{aligned}(\mathbf{I} - \hat{\mathbf{B}}^{(i)})^{-1} &= \begin{bmatrix} \left(\mathbf{I} - \tilde{\mathbf{C}}_i\right)^{-1} & -\left(\mathbf{I} - \tilde{\mathbf{C}}_i\right)^{-1}(-\hat{\mathbf{B}}_{i,:})(\mathbf{I} - \mathbf{B}_{\backslash i})^{-1} \\ \mathbf{Y}_1 & \mathbf{Y}_2 \end{bmatrix} \\ &= \begin{bmatrix} \left(\mathbf{I} - \tilde{\mathbf{C}}_i\right)^{-1} & \left(\mathbf{I} - \tilde{\mathbf{C}}_i\right)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1} \\ \mathbf{Y}_1 & \mathbf{Y}_2 \end{bmatrix},\end{aligned}$$

where $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are matrices which is not of our further interest. Multiplying with $\mathbf{I} - \mathbf{\Gamma}^{(i)}$, we obtain:

$$(\mathbf{I} - \hat{\mathbf{B}}^{(i)})^{-1}(\mathbf{I} - \mathbf{\Gamma}^{(i)}) = \begin{bmatrix} (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{I} - \mathbf{\Gamma}_i) & \mathbf{F}^{(i)} \\ \mathbf{Y}_1 & \mathbf{Y}_2 \end{bmatrix},$$

where:

$$\mathbf{F}^{(i)} := (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I} - \mathbf{\Gamma}_{\backslash i})$$

Now, we aim to transfer above result to $(\mathbf{I} - \mathbf{B})^{-1}(\mathbf{I} - \mathbf{\Gamma})$. First step toward this end is to notice that the identity (9.27)gives:

$$\begin{aligned}\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\hat{\mathbf{B}}^{(i)})^{-1}\mathbf{Q}_{i\leftrightarrow 1} &= \sum_{k=0}^{\infty}\mathbf{Q}_{i\leftrightarrow 1}\hat{\mathbf{B}}^{(i),k}\mathbf{Q}_{i\leftrightarrow 1} = \sum_{k=0}^{\infty}\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{Q}_{i\leftrightarrow 1}\mathbf{B}\mathbf{Q}_{i\leftrightarrow 1})^{k}\mathbf{Q}_{i\leftrightarrow 1}\\ &= \sum_{k=0}^{\infty}\mathbf{B}\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{Q}_{i\leftrightarrow 1}\mathbf{B}\mathbf{Q}_{i\leftrightarrow 1})^{k-1}\mathbf{Q}_{i\leftrightarrow 1}\\ &= \sum_{k=0}^{\infty}\mathbf{B}^2\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{Q}_{i\leftrightarrow 1}\mathbf{B}\mathbf{Q}_{i\leftrightarrow 1})^{k-2}\mathbf{Q}_{i\leftrightarrow 1}\\ &= \ldots = \sum_{k=0}^{\infty}\mathbf{B}^k\mathbf{Q}_{i\leftrightarrow 1}\mathbf{Q}_{i\leftrightarrow 1} = \sum_{k=0}^{\infty}\mathbf{B}^k = (\mathbf{I}-\mathbf{B})^{-1}.\end{aligned}$$

Consequently:

$$\begin{aligned}\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\hat{\mathbf{B}}^{(i)})^{-1}(\mathbf{I}-\hat{\mathbf{\Gamma}}^{(i)})\mathbf{Q}_{i\leftrightarrow 1} &= \mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\hat{\mathbf{B}}^{(i)})^{-1}(\mathbf{I}-\mathbf{Q}_{i\leftrightarrow 1}\mathbf{\Gamma}\mathbf{Q}_{i\leftrightarrow 1})\mathbf{Q}_{i\leftrightarrow 1}\\ &= \mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\hat{\mathbf{B}}^{(i)})^{-1}\mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{Q}_{i\leftrightarrow 1}\mathbf{Q}_{i\leftrightarrow 1} \qquad (9.28)\\ &= (\mathbf{I}-\mathbf{B})^{-1}(\mathbf{I}-\mathbf{\Gamma}).\end{aligned}$$

So since $\mathbf{Q}_{i\leftrightarrow 1}(\cdot)\mathbf{Q}_{i\leftrightarrow 1}$ corresponds to the interchange of the first and i-th block rows of $(\cdot)$, and the first and i-th block columns of $(\cdot)$, it follows from the identity (9.28), that the i-th block row of $(\mathbf{I}-\mathbf{B})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ results by block column permutation of the block row matrix $\left[(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{I}-\mathbf{\Gamma}_i),\mathbf{F}^{(i)}\right]$ where the i-th block column of the i-th block row of $(\mathbf{I}-\mathbf{B})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ is $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)$. ■

Now, we continue to exploit 9.17 by giving the following further auxiliary statements

Lemma 9.18: *We have the following statements:*

1. $(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})\geqslant 0$
2. $\left[(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})\right]_{(i):}\mathbf{1} = (\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{1}$

Proof: Lemma 9.17 asserts that the block diagonal of $(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ coincides with the block diagonal of $(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})$. Since the latter matrix is a block diagonal matrix, it follows that $(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ is a counter block diagonal matrix having the block row permutation of $\mathbf{F}^{(i)} := (\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})$, $i\in[M]$ as off-diagonal entries. By the obvious fact that $\mathbf{F}^{(i)}\geqslant 0$ for any $i\in[M]$, we obtain the first statement.

Now, to proof the second statement, notice again by Lemma 9.17, that the off block diagonal on the i-th block row of $(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ is a block column permutation of $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})$. Since $(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ is a counter block diagonal matrix (see last paragraph) and the operation $(\cdot)\mathbf{1}$ is invariant of column permutation of $(\cdot)$, the second statement follows. ■

The expression on the R.H.S. of the identity given in above Lemma is cumbersome. By this reason, we show the following:

Lemma 9.19:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}]\mathbf{1} = (\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{\Gamma}_i-\tilde{\mathbf{C}}_i)\mathbf{1}.$$

Proof: We have from Lemma 9.13:

$$\begin{aligned}
&\mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}]\mathbf{1}\\
&=\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1}-\mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}]\mathbf{1}\\
&=(\mathbf{\Gamma}_i-\mathbf{\Gamma}_i\mathbf{A}_{ii}-\mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i})\mathbf{1}\\
&=(\mathbf{\Gamma}_i-\tilde{\mathbf{C}}_i)\mathbf{1}
\end{aligned}$$

Consequently:

$$\begin{aligned}
&(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}]^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}]\mathbf{1}\\
&=(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{\Gamma}_i-\tilde{\mathbf{C}}_i)
\end{aligned}\qquad\blacksquare$$

Now, we are ready to give a proof of Theorem 9.16:

Proof (Proof of Theorem 9.16): Since row sum is invariant of column permutation, (9.23) follows from Lemma 9.17. Now, setting the identity given in Lemma 9.19 into (9.23), we obtain (9.24). Now to see the last remaining identity, notice that by the first fact in Lemma 9.18, we have that $\mathbf{h}=\Big[(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})\Big]\mathbf{1}$. This identity and Lemma 9.3 asserts:

$$\mathbf{h}=(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}=\mathbf{1}-(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}.$$

Since $(\mathbf{I}-\tilde{\mathbf{C}})^{-1}(\mathbf{I}-\mathbf{\Gamma})$ is a block diagonal matrix with the matrices $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)$, $i\in[N]$, on the diagonals, we obtain the desired statement. $\blacksquare$

9.5. Limit Behaviour of the Mesoscopic Stability of Opinion Dynamics in face of Informational Bias

Based on our findings, we aim to analyze to what extent does the informational bias effects in clusterness of the population's opinions. In doing so, we consider to limit cases:

- First, the case of almost stubborn agents, where the agents susceptibility constant is almost 0, or equivalently, the bias constant is almost 1
- second, the case where the agents are subject to negligible amount of informational

bias, i.e. the case where the agents susceptibility constant is almost 1, or equivalently, the bias constant is almost 0.

The tool for this task is the measure $\mathbf{h}$, which quantifies the worst-case asymptotic error of the approximation of the opinion dynamic by means of the cluster-centric dynamic formed by the substochastic complementation of the communication matrix.

9.5.1. Behaviour for $\mathbf{\Gamma}_i \approx 0$

For a cluster $i \in [M]$, we first investigate the behaviour of $\mathbf{h}_i$ in case that all agents are very much affected by informational bias. Our result is the following:

Proposition 9.20: *It holds:*

$$\mathbf{h}_i \leqslant \mathcal{O}(\|\mathbf{\Gamma}_i\|).$$

Proof: We first show:

$$\mathbf{h}_i \leqslant \|\mathbf{\Gamma}_i\|_\infty + \|\mathbf{I} - \mathbf{\Gamma}_i\|_\infty \|\tilde{\mathbf{C}}_i\|_\infty. \tag{9.29}$$

From the identity (9.25), we have:

$$\mathbf{h}_i = \mathbf{1} - (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1} = \mathbf{1} - \sum_{k=0}^{\infty} \tilde{\mathbf{C}}_i^k (\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1}.$$

Since $(\mathbf{I} - \mathbf{\Gamma}_i)$ and $\tilde{\mathbf{C}}_i^k$, for any $k \in \mathbb{N}$, are non-negative it yields:

$$\sum_{k=0}^{\infty} \tilde{\mathbf{C}}_i^k (\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1} \geqslant (\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1} + \tilde{\mathbf{C}}_i(\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1}.$$

Consequently it holds:

$$\mathbf{h}_i \leqslant \mathbf{\Gamma}_i \mathbf{1} + \tilde{\mathbf{C}}_i(\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1},$$

from which we have (9.29) by straightforward computations.

Now, in order to show this proposition, it remains to provide the estimation $\tilde{\mathbf{C}}_i\mathbf{1} \leqslant \mathcal{O}(\|\mathbf{\Gamma}_i\|)$. This is already done in (9.14), where we have seen that $\tilde{\mathbf{C}}_i\mathbf{1} \leqslant \mathbf{\Gamma}_i\mathbf{1}$. ■

So from above result it follows that $\mathbf{h}_i$ is small if the susceptibility constant of the agents in cluster i is small irrespective of the degree of the perturbation of other agents by informational bias. Furthermore, we are able to specify that the decrease of $\mathbf{h}_i$ respective to $\mathbf{\Gamma}_i$ is at worst linear. We can refine above results by providing the second-order term of the decrease:

Theorem 9.21: *It holds:*

$$\mathbf{h}_i = \mathbf{\Gamma}_i \mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i})\mathbf{1} + \mathcal{O}(\|\mathbf{\Gamma}_i\|^2 \|\mathbf{A}_{i:}\|).$$

Proof: We have:

$$\tilde{\mathbf{C}}_i = \mathbf{\Gamma}_i \mathbf{A}_{ii} + \mathbf{\Gamma}_i \mathbf{A}_{i:} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i} \mathbf{A}_{\setminus i})^{-1} \mathbf{\Gamma}_{\setminus i} \mathbf{A}_{:i} = \mathcal{O}(\|\mathbf{\Gamma}_i\|).$$

This observation gives:

$$(\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1} = \mathbf{I} + \mathcal{O}(\|\tilde{\mathbf{C}}_i\|) = \mathbf{I} + \mathcal{O}(\|\mathbf{\Gamma}_i\|). \tag{9.30}$$

Finally we obtain the desired result:

$$\begin{aligned}
\mathbf{h}_i &= (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1} \mathbf{\Gamma}_i \mathbf{A}_{i:} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i} \mathbf{A}_{\setminus i})^{-1} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i}) \mathbf{1} \\
&= [\mathbf{I} + \mathcal{O}(\|\mathbf{\Gamma}_i\|)]\, \mathbf{\Gamma}_i \mathbf{A}_{i:} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i} \mathbf{A}_{\setminus i})^{-1} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i}) \mathbf{1} \\
&= \mathbf{\Gamma}_i \mathbf{A}_{i:} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i} \mathbf{A}_{\setminus i})^{-1} (\mathbf{I} - \mathbf{\Gamma}_{\setminus i}) \mathbf{1} + \mathcal{O}(\|\mathbf{\Gamma}_i\|^2 \|\mathbf{A}_{i:}\|)
\end{aligned}$$

■

From Proposition 9.20, the question raises gives whether $\mathbf{h}_i$ decreases with increasing influence of the informational bias to the agents' opinion. Intuitively, the answer should be positive since, we expect that by the increasing bias, the amount of information exchange between the agents decreases. In the following, we provide a partial answer to this question:

Proposition 9.22: *Suppose that:*

$$\mathbf{\Gamma} = \gamma \mathbf{I}, \quad \gamma > 0.$$

Then it holds that $\mathbf{h}$ *is (entrywise) non-decreasing for* $\gamma \in (0.0.5)$

Proof: We have:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\gamma}(1-\gamma)\left(\mathbf{I} - \tilde{\mathbf{C}}_i(\gamma)\right)^{-1} &= \frac{\mathrm{d}}{\mathrm{d}\gamma}(1-\gamma)\sum_{n=0}^{\infty} \tilde{\mathbf{C}}_i(\gamma)^k = -\sum_{k=0}^{\infty} \tilde{\mathbf{C}}_i(\gamma)^k + (1-\gamma)\sum_{n=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}\gamma} \tilde{\mathbf{C}}_i(\gamma)^k \\
&= \sum_{n=0}^{\infty} \left[(1-\gamma)\frac{\mathrm{d}}{\mathrm{d}\gamma} \tilde{\mathbf{C}}_i(\gamma)^k - \tilde{\mathbf{C}}_i(\gamma)^k\right]
\end{aligned}$$

Now, for any $n \in \mathbb{N}$, it holds:

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} \tilde{\mathbf{C}}_i(\gamma)^n = \sum_{k=1}^{n} \tilde{\mathbf{C}}_i(\gamma)^{k-1} \left[\frac{\mathrm{d}}{\mathrm{d}\gamma} \tilde{\mathbf{C}}_i(\gamma)\right] \tilde{\mathbf{C}}_i(\gamma)^{n-k},$$

and:

$$\tilde{\mathbf{C}}_i(\gamma)^n = \sum_{k=1}^{n} \tilde{\mathbf{C}}_i(\gamma)^{k-1} \frac{\tilde{\mathbf{C}}_i(\gamma)}{n} \tilde{\mathbf{C}}_i(\gamma)^{n-k}.$$

Consequently:

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}(1-\gamma)\left(\mathbf{I}-\tilde{\mathbf{C}}_i(\gamma)\right)^{-1} = \sum_{n=0}^{\infty}\sum_{k=1}^{n}\tilde{\mathbf{C}}_i(\gamma)^{k-1}\left[(1-\gamma)\frac{\mathrm{d}}{\mathrm{d}\gamma}\tilde{\mathbf{C}}_i(\gamma)-\frac{1}{n}\tilde{\mathbf{C}}_i(\gamma)\right]\tilde{\mathbf{C}}_i(\gamma)^{n-k}.$$

Next, we have:

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\tilde{\mathbf{C}}_i(\gamma) = \mathbf{A}_{ii} + 2\gamma\mathbf{A}_{i:}(\mathbf{I}-\gamma\mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i} + \gamma^2\mathbf{A}_{i:}\frac{\mathrm{d}}{\mathrm{d}\gamma}(\mathbf{I}-\gamma\mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i},$$

and consequently:

$$\gamma\frac{\mathrm{d}}{\mathrm{d}\gamma}\tilde{\mathbf{C}}_i(\gamma) = \tilde{\mathbf{C}}_i(\gamma) + \mathbf{R}(\gamma),$$

where:

$$\mathbf{R}(\gamma) := \gamma^2\mathbf{A}_{i:}(\mathbf{I}-\gamma\mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i} + \gamma^3\mathbf{A}_{i:}\frac{\mathrm{d}}{\mathrm{d}\gamma}(\mathbf{I}-\gamma\mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{\cdot i:}$$

Thus:

$$(1-\gamma)\frac{\mathrm{d}}{\mathrm{d}\gamma}\tilde{\mathbf{C}}_i(\gamma) - \frac{1}{n}\tilde{\mathbf{C}}_i(\gamma) = \left(\frac{1-\gamma}{\gamma}-\frac{1}{n}\right)\tilde{\mathbf{C}}_i + \frac{1-\gamma}{\gamma}\mathbf{R}(\gamma).$$

If $\gamma \leqslant 0.5$, then $(1-\gamma)/\gamma \geqslant 1$. Consequently since $\tilde{\mathbf{C}}_i(\gamma), \mathbf{R}(\gamma) \geqslant 0$, we have:

$$\begin{aligned}
&(1-\gamma)\frac{\mathrm{d}}{\mathrm{d}\gamma}\tilde{\mathbf{C}}_i(\gamma) - \frac{1}{n}\tilde{\mathbf{C}}_i(\gamma) = \left(\frac{1-\gamma}{\gamma}-\frac{1}{n}\right)\tilde{\mathbf{C}}_i + \frac{1-\gamma}{\gamma}\mathbf{R}(\gamma)\\
&\geqslant \underbrace{\left(\frac{1-\gamma}{\gamma}-1\right)}_{\geqslant 0}\tilde{\mathbf{C}}_i + \frac{1-\gamma}{\gamma}\mathbf{R}(\gamma) \geqslant \frac{1-\gamma}{\gamma}\mathbf{R}(\gamma) \geqslant 0
\end{aligned}$$

This shows that $\mathbf{I}-(1-\gamma)(\mathbf{I}-\tilde{\mathbf{C}}(\gamma))^{-1}$ is (entrywise) monotonically increasing. Finally, we obtain from (9.25) the desired statement. ■

As shown in Proposition 9.20, if the agents in a cluster $i \in [N]$ are very much influenced by informational bias, then irrespective of how much out-cluster agents are influenced by informational bias $\mathbf{h}_i$ is small. Now, the following question arises: how does $\mathbf{h}_i$ behaves if the out-cluster agents are object to a high degree of perturbation by informational bias. To answer this question, we first need the following identity which is a direct consequence of the Woodbury identity:

Lemma 9.23: *It holds:*

$$(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1} = (\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\left[\mathbf{I}+\mathbf{\Gamma}_i\mathbf{A}_{i:}\left(\mathbf{I}-\tilde{\mathbf{C}}_i^{\circlearrowleft}\right)^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\right], \tag{9.31}$$

where:

$$\tilde{\mathbf{C}}_i^{\circlearrowleft} = \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i} + \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}.$$

We continue by providing an estimate for $\tilde{\mathbf{C}}_i^{\circlearrowleft}$:

Lemma 9.24:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}=(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\left[\mathbf{I}+\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\right]+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2\|\mathbf{A}_{:i}\|)$$

Proof: We aim estimate $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}$ via the identity (9.31). First notice that:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i^{\circlearrowright})^{-1}=\mathbf{I}+\tilde{\mathbf{C}}_i^{\circlearrowright}+\mathcal{O}(\|\tilde{\mathbf{C}}_i^{\circlearrowright}\|^2). \tag{9.32}$$

Moreover:

$$\begin{aligned}\tilde{\mathbf{C}}_i^{\circlearrowright}&=\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}+\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_{ii}\mathbf{A}_{ii})^{-1}\mathbf{A}_{i:}=\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|)\\&=-\mathbf{\Gamma}_{\setminus i}+\mathbf{\Gamma}_{\setminus i}(\mathbf{I}-\mathbf{A}_{\setminus i})+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|)\leqslant-\mathbf{\Gamma}_{\setminus i}+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{I}-\mathbf{A}_{\setminus i}\|)+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|)\\&=-\mathbf{\Gamma}_{\setminus i}+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|),\end{aligned}$$

and particularly:

$$\tilde{\mathbf{C}}_i^{\circlearrowright}=\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|).$$

Setting both previous estimates into (9.32) , it yields:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i^{\circlearrowright})^{-1}=(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|)+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2).$$

From above computation, we aim to provide an estimate for $(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}$ having the representation (9.31). Toward this direction, we estimate:

$$\begin{aligned}&\mathbf{\Gamma}_i\mathbf{A}_{i:}\left(\mathbf{I}-\tilde{\mathbf{C}}_i^{\circlearrowright}\right)^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\\&=\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|)\left(\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|)+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2)\right)\\&=\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2\|\mathbf{A}_{:i}\|).\end{aligned}$$

Consequently:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}=(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\left[\mathbf{I}+\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\right]+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2\|\mathbf{A}_{:i}\|)\quad\blacksquare$$

Now we can provide the answer of the question given in the beginning of this subsection:

Theorem 9.25: *Let $\mathbf{\Gamma}_i$, $\mathbf{A}_{ii}$ be fix. It holds:*

$$\begin{aligned}\mathbf{h}_i&=(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1}\\&\quad-(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1}\\&\quad+\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2\|\mathbf{A}_{:i}\|),\end{aligned}$$

and:

Proof: Now we have:

$$(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i)\mathbf{1} = \mathbf{\Gamma}_i\mathbf{A}_{i:}\left(\mathbf{I} - (\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\right) = \mathbf{\Gamma}_i\mathbf{A}_{i:} + \mathcal{O}\left(\frac{\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|}{1 - \|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{\setminus i}\|}\right),$$

and since $\|\mathbf{A}_{:i}\|_\infty = 1 - \|\mathbf{A}_{\setminus i}\|_\infty$, we have:

$$(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i)\mathbf{1} = \mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} + \mathcal{O}\left(\frac{\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|}{1 - \|\mathbf{\Gamma}_{\setminus i}\| + \|\mathbf{A}_{:i}\|}\right) \leqslant \mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} + \mathcal{O}\left(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|\right).$$

Finally, we have by (9.23) and Lemma 9.24:

$$\mathbf{h}_i = (\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\left[\mathbf{I} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\right]\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} + \mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|\|\mathbf{A}_{:i}\|) \quad \blacksquare$$

Above Theorem asserts in particular that if the out-cluster agents are almost fully biased, then $\mathbf{h}_i$ depends only on the structure and behaviour of cluster i. This is quite surprising, since the main believe is that the opinion dynamic is determined by the most biased agents.

We close this subsection by providing a summary of our results in Table 9.1.

Table 9.1.: An Overview of Asymptotic behaviour of $\mathbf{h}_i$

Asymptote	Error Scale	$\mathbf{h}_i \approx$
$\mathbf{\Gamma}_i \to 0$	$\mathcal{O}(\|\mathbf{\Gamma}_i\|)$	0
	$\mathcal{O}(\|\mathbf{\Gamma}_i\|^2)$	$\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i})\mathbf{1}$
$\mathbf{\Gamma}_{\setminus i} \to 0$	$\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|)$	$(\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1}$
	$\mathcal{O}(\|\mathbf{\Gamma}_{\setminus i}\|^2)$	$(\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1}$ $-(\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i})\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}(\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}(\mathbf{I} - \mathbf{\Gamma}_i)\mathbf{1}$
$\mathbf{A}_{:i} \to 0$	$\mathcal{O}(\|\mathbf{A}_{:i}\|)$	$(\mathbf{I} - \mathbf{\Gamma}_i\mathbf{A}_{ii})^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}$
$\mathbf{A}_{i:} \to 0$	$\mathcal{O}(\|\mathbf{A}_{i:}\|)$	0

9.5.2. Behaviour for $\mathbf{\Gamma} \approx \mathbf{I}$

Now, we investigate the quantity $\mathbf{h}$ in case that the degree of influence of informational bias is small in comparison to the opinion mixing.

We start with the case where for a given cluster i, either the in-cluster or out-cluster agents are not biased by extrinsic information:

Lemma 9.26: *If $\mathbf{\Gamma}_i = \mathbf{I}$, then for any $\mathbf{\Gamma}_{\setminus i}$:*

$$\mathbf{h}_i = \mathbf{1}.$$

Moreover if $\mathbf{\Gamma}_{\setminus i} = \mathbf{I}$*, then for any* $\mathbf{\Gamma}_i$*:*

$$\mathbf{h}_i = 0.$$

Proof: The first statement is shown by the following elementary computation:

$$\mathbf{h}_i = (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i)\mathbf{1} = (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{I} - \tilde{\mathbf{C}}_i)\mathbf{1} = \mathbf{1}.$$

The second statement follows since:

$$\begin{aligned}
\tilde{\mathbf{C}}_i\mathbf{1} &= \mathbf{\Gamma}_i\mathbf{A}_{ii}\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\mathbf{1} = \mathbf{\Gamma}_i\mathbf{A}_{ii}\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i}\mathbf{1} \\
&= \mathbf{\Gamma}_i\mathbf{A}_{ii}\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I} - \mathbf{A}_{\setminus i})^{-1}(\mathbf{I} - \mathbf{A}_{\setminus i})\mathbf{1} = \mathbf{\Gamma}_i\mathbf{A}_{ii}\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}\mathbf{1} \\
&= \mathbf{\Gamma}_i(\mathbf{A}_{ii} + \mathbf{A}_{ii})\mathbf{1} = \mathbf{\Gamma}_i\mathbf{1},
\end{aligned}$$

which yields $(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i)\mathbf{1} = 0$ and finally $\mathbf{h}_i = (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i)\mathbf{1} = 0$ ■

So above Lemma asserts that if the in-cluster agents are not subject to informational bias then for any negligible disturbance of the out-cluster agents by informational bias, it holds that $\mathbf{h}_i = \mathbf{1}$. In case that the out-cluster agent are not subject to informational bias, we have a different situation: For any negligible disturbance of the in-cluster agents by informational bias, it holds that $\mathbf{h}_i = 0$.

Above observation gives the hint that the behaviour of $\mathbf{h}_i$ in case that the agents are subject to negligible disturbance of informational bias cannot be predicted in a uniform way. The following proposition corfirms in some sense this guess by showing that the behaviour of $\mathbf{h}_i$ can in general be arbitrary:

Proposition 9.27: *Suppose that there exists* $p_i, p_{\setminus i} > 0$ *s.t.:*

$$\mathbf{A}_{ii}\mathbf{1} = p_i\mathbf{1} \quad \textit{and} \quad \mathbf{A}_{\setminus i}\mathbf{1} = p_{\setminus i}\mathbf{1}. \tag{9.33}$$

For any $\alpha \in [0, 1]$*, there exists a sequence* $(\mathbf{\Gamma}(n))_{n\in\mathbb{N}}$ *converging to* $\mathbf{I}$*, such that:*

$$\lim_{n\to\infty} \mathbf{h}_i(\mathbf{\Gamma}(n)) = \alpha\mathbf{1}$$

The proof of above statement is based on the following auxiliary statement:

Lemma 9.28: *Suppose that* (9.33) *holds. Then for* $\mathbf{\Gamma} = \mathrm{diag}(\gamma_i\mathbf{I}, \gamma_{\setminus i}\mathbf{I})$*:*

$$\mathbf{h}_i = \frac{1}{1 + \psi^{(i)}}, \quad \textit{where} \quad \psi^{(i)} := \frac{\frac{\gamma_{\setminus i}}{1-\gamma_{\setminus i}}}{\frac{\gamma_i}{1-\gamma_i}} \frac{\frac{1}{\gamma_{\setminus i}} - p_{\setminus i}}{1 - p_i}$$

Proof: We have:

$$\begin{aligned}\tilde{\mathbf{C}}_i\mathbf{1} &= \left(\mathbf{\Gamma}_i\mathbf{A}_{ii} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{:i}\right)\mathbf{1} = \left(\gamma_i\mathbf{A}_{ii} + \gamma_i\gamma_{\setminus i}\mathbf{A}_{i:}(\mathbf{I}-\gamma_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{A}_{:i}\right)\mathbf{1}\\ &= \gamma_i p_i\mathbf{1} + q_{\setminus i}\gamma_i\gamma_{\setminus i}\mathbf{A}_{i:}(\mathbf{I}-\gamma_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{1},\end{aligned}$$

where $q_i := 1 - p_i$ satisfying:

$$q_i\mathbf{1} = (1-p_i)\mathbf{1} = (\mathbf{I}-\mathbf{A}_{ii})\mathbf{1} = \mathbf{A}_{i:}\mathbf{1},$$

and $q_{\setminus i} := 1 - p_{\setminus i}$ satisfying:

$$q_{\setminus i}\mathbf{1} = (1-p_{\setminus i})\mathbf{1} = (\mathbf{I}-\mathbf{A}_{\setminus i})\mathbf{1} = \mathbf{A}_{:i}\mathbf{1}.$$

Now, we have:

$$(\mathbf{I}-\gamma_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{1} = \sum_{k=0}^{\infty}\gamma_{\setminus i}^k\mathbf{A}_{\setminus i}^k\mathbf{1} = \sum_{k=0}^{\infty}\gamma_{\setminus i}^k p_{\setminus i}^k\mathbf{1} = \frac{1}{1-\gamma_{\setminus i}p_{\setminus i}}\mathbf{1}.$$

Setting above identity into (71), it holds:

$$\begin{aligned}\tilde{\mathbf{C}}_i\mathbf{1} &= \gamma_i p_i\mathbf{1} + q_{\setminus i}\gamma_i\gamma_{\setminus i}\mathbf{A}_{i:}(\mathbf{I}-\gamma_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\mathbf{1} = \gamma_i p_i\mathbf{1} + \frac{q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\mathbf{A}_{i:}\mathbf{1} = \gamma_i p_i\mathbf{1} + \frac{q_i q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\mathbf{1}\\ &= \left(\gamma_i p_i + \frac{q_i q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\right)\mathbf{1}.\end{aligned}$$

Consequently, we have:

$$(\mathbf{I}-\tilde{\mathbf{C}}_i)^{-1}\mathbf{1} = \sum_{k=0}^{\infty}\tilde{\mathbf{C}}_i^k\mathbf{1} = \sum_{k=0}^{\infty}\left(\gamma_i p_i + \frac{q_i q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\right)^k\mathbf{1} = \frac{1}{1-\left(\gamma_i p_i + \frac{q_i q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\right)}\mathbf{1}.$$

Now, we have:

$$\left(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i\right)\mathbf{1} = \left(\gamma_i - \gamma_i p_i - \frac{q_i q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\right)\mathbf{1} = \left(\gamma_i q_i - \frac{q_i q_{\setminus i}\gamma_i\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\right)\mathbf{1} = \gamma_i q_i\left(1 - \frac{q_{\setminus i}\gamma_{\setminus i}}{1-\gamma_{\setminus i}p_{\setminus i}}\right)\mathbf{1}$$

Finally, we have:

$$\begin{aligned}
\mathbf{h}_i(\gamma_i, \gamma_{\backslash i}) = (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{\Gamma}_i - \tilde{\mathbf{C}}_i)\mathbf{1} &= \frac{\gamma_i q_i \left(1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}{1 - \left(\gamma_i p_i + \frac{q_i q_{\backslash i}\gamma_i\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}\mathbf{1} = \frac{\gamma_i q_i \left(1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}{1 - \left(\gamma_i p_i + \frac{q_i q_{\backslash i}\gamma_i\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}\mathbf{1} \\
&= \frac{\gamma_i q_i \left(1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}{1 - \gamma_i(1-q_i) - \frac{q_i q_{\backslash i}\gamma_i\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}}\mathbf{1} = \frac{\gamma_i q_i \left(1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}{(1-\gamma_i) + \gamma_i q_i \left(1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}\mathbf{1} \\
&= \frac{1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}}{\frac{1-\gamma_i}{\gamma_i q_i} + \left(1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}}\right)}\mathbf{1} = \frac{1}{1 + \frac{\psi_i^{(i)}(\gamma_i)}{\psi_{\backslash i}^{(i)}(\gamma_{\backslash i})}}\mathbf{1},
\end{aligned} \tag{9.34}$$

where:

$$\psi_i^{(i)} := \frac{1-\gamma_i}{\gamma_i q_i} = \frac{1-\gamma_i}{\gamma_i(1-p_i)} \quad \text{and} \quad \psi_{\backslash i}^{(i)} := 1 - \frac{q_{\backslash i}\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}} = \frac{1-\gamma_{\backslash i}}{1-\gamma_{\backslash i}p_{\backslash i}} = \frac{\frac{1-\gamma_{\backslash i}}{\gamma_{\backslash i}}}{\frac{1}{\gamma_{\backslash i}} - p_{\backslash i}}.$$

We obtain finally the desired result by noticing that $\frac{\psi_i^{(i)}}{\psi_{\backslash i}^{(i)}} = \psi^{(i)}$. ■

Now, we are ready to proof Proposition 9.27

Proof (Proof of Proposition 9.27): We only show the statement for $i = 1$. The desired result can be established from this case by permutation argument. Define:

$$\beta := \frac{1-\alpha}{\alpha} \text{and} \quad \tilde{\beta} = \beta\frac{1-p_1}{1-p_{\backslash 1}}.$$

Let $N \in \mathbb{N}$ be large enough such that:

$$\frac{\tilde{\beta}}{N} \leqslant 1.$$

We define:

$$\mathbf{\Gamma}(n) := \mathrm{diag}(\gamma_1(n)\mathbf{I}, \gamma_{\backslash 1}(n)\mathbf{I})$$

with:

$$1 - \gamma_1(n) := \frac{\tilde{\beta}}{n},\ n \geqslant N, \qquad \text{and} \qquad 1 - \gamma_{\backslash i}(n) := \frac{1}{n},\ n \in \mathbb{N}.$$

We have $\lim_{n\to\infty} \gamma_1(n) = 1$ and $\lim_{n\to\infty} \gamma_{\backslash i}(n) = 1$. Consequently:

$$\lim_{n\to\infty} \psi^{(1)} = \tilde{\beta}\frac{1-p_{\backslash 1}}{1-p_1} = \beta.$$

This yields as desired:

$$\lim_{n\to\infty} \mathbf{h}_1 = \lim_{n\to\infty} \frac{1}{1+\psi^{(1)}} = \frac{1}{1+\beta} = \alpha.$$ ■

Nevertheless, we can specify the following:

Lemma 9.29: *Suppose that* $\mathbf{A}$ *is irreducible. Let* $(\boldsymbol{\Gamma}(n))_n$ *be a sequence converging to the identity. Then the set of limit points of* $(\mathbf{h}_i(\boldsymbol{\Gamma}(n)))_n$ *is contained in the set:*

$$\mathcal{C} := \{\alpha\mathbf{1} : \ \alpha \in [0,1]\}.$$

Proof: Let $\overline{\mathbf{h}_i}$ be a limit point of $(\mathbf{h}_i(\boldsymbol{\Gamma}(n)))_n$, then there exists a subsequence $(\mathbf{h}_i(\boldsymbol{\Gamma}(n_k)))_k$ of $(\mathbf{h}_i(\boldsymbol{\Gamma}(n)))_n$ converging to $\overline{\mathbf{h}_i}$.

Now, we have:

$$\begin{aligned}(\mathbf{I} - \tilde{\mathbf{C}}_i(\boldsymbol{\Gamma}(n_k)))\mathbf{h}_i(\boldsymbol{\Gamma}(n_k)) &= (\mathbf{I} - \tilde{\mathbf{C}}_i(\boldsymbol{\Gamma}(n_k)))(\mathbf{I} - \tilde{\mathbf{C}}_i(\boldsymbol{\Gamma}(n_k)))^{-1}(\boldsymbol{\Gamma}(n_k) - \tilde{\mathbf{C}}_i(\boldsymbol{\Gamma}(n_k)))\mathbf{1} \\ &= (\boldsymbol{\Gamma}(n_k) - \tilde{\mathbf{C}}_i(\boldsymbol{\Gamma}(n_k)))\mathbf{1}.\end{aligned}$$

Since $\tilde{\mathbf{C}}_i(\mathbf{I}) = \mathbf{C}_i$, we obtain by letting $k \to \infty$ on both sides of above identity

$$(\mathbf{I} - \mathbf{C}_i)\overline{\mathbf{h}_i} = \lim_{k\to\infty}(\mathbf{I} - \tilde{\mathbf{C}}_i)\mathbf{h}_i = \lim_{k\to\infty}(\boldsymbol{\Gamma}(n_k) - \tilde{\mathbf{C}}_i(\boldsymbol{\Gamma}(n_k)))\mathbf{1} = (\mathbf{I} - \mathbf{C}_i)\mathbf{1} = 0,$$

so that:

$$\mathbf{C}_i\overline{\mathbf{h}_i} = \overline{\mathbf{h}_i}.$$

Since $\mathbf{C}_i$ is a irreducible stochastic matrix, which is a consequence of the irreducibility assumption of $\mathbf{A}$ and Lemma 9.8, we have by Perron-Frobenius Theorem the desired statement. ∎

Above result says that in case the influence of informational bias to the agents is negligible, the cluster centric approximation error remains the same. However, it does not specify the corresponding quantity.

9.5.3. Detailed Investigation for $\boldsymbol{\Gamma} \approx \mathbf{I}$

Our aim in this subsection is to specify the behaviour of $\mathbf{h}_i$ in case that $\boldsymbol{\Gamma} \approx \mathbf{I}$. In the following we use the following notations. $\gamma_{\max}^{(i)}$ (resp. $\gamma_{\min}^{(i)}$) denotes the largest (resp. the smallest) susceptibility constant of the agents in cluster i. Furthermore, we denote the largest (resp. the smallest) susceptibility constant of the agents out of cluster i by $\gamma_{\max}^{(\backslash i)}$ (resp. $\gamma_{\min}^{(\backslash i)}$). For a matrix $\mathbf{B}$, $\|\mathbf{B}\|_{\min}$ denotes the row sum of $\mathbf{B}$ with the smallest absolute value.

Define:

$$\mathbf{G}_i := (\mathbf{I} - \tilde{\mathbf{C}}_i)^{-1}(\mathbf{I} - \boldsymbol{\Gamma}_i). \tag{9.35}$$

Notice that by (9.25), it holds:

$$\mathbf{h}_i = \mathbf{1} - \mathbf{G}_i\mathbf{1}. \tag{9.36}$$

Therefore, we can work with $\mathbf{G}_i$ instead of directly with $\mathbf{h}_i$. However, working with $\mathbf{G}_i$

is not handy since this terms involves the infinite sum of $\tilde{\mathbf{C}}_i$, which is itself an involved term. By this reason, we decide to work with:

$$\mathbf{G}_i^{-1} = (\mathbf{I} - \mathbf{\Gamma}_i)^{-1}(\mathbf{I} - \tilde{\mathbf{C}}_i).$$

Our candidate for the limit matrix is the following rank one matrix:

$$\overline{\mathbf{G}}_i := \frac{\mathbf{1}\eta^{\mathrm{T}}}{\mathrm{c}_i},$$

where $\eta \in \mathbb{R}^{K_i}$ and c_i set later and are possibly dependent on $\mathbf{\Gamma}$. From Woodbury-identity, it follows:

$$(\mathbf{G}_i - \overline{\mathbf{G}}_i)^{-1} = \mathbf{G}_i^{-1} + \frac{(\mathbf{G}_i^{-1}\mathbf{1}\eta_i^{\mathrm{T}}\mathbf{G}_i^{-1})/\mathrm{c}_i}{1 - (\eta_i^{\mathrm{T}}\mathbf{G}_i^{-1}\mathbf{1})/\mathrm{c}_i},$$

and consequently:

$$(\mathbf{G}_i - \overline{\mathbf{G}}_i)^{-1}\mathbf{1} = \mathbf{G}_i^{-1}\mathbf{1}\left(1 + \frac{\frac{\mathrm{d}_i}{\mathrm{c}_i}}{1 - \frac{\mathrm{d}_i}{\mathrm{c}_i}}\right), \tag{9.37}$$

where:

$$\mathrm{d}_i := \eta_i^{\mathrm{T}}\mathbf{G}_i^{-1}\mathbf{1}. \tag{9.38}$$

Now, our strategy is to choose η_i and c_i such that $\mathrm{d}_i/\mathrm{c}_i \approx 1$ whenever $\mathbf{\Gamma} \approx \mathbf{I}$. Moreover, $\mathbf{G}_i^{-1}\mathbf{1}$ is not small, the latter and (9.37) implies that $(\mathbf{G}_i - \overline{\mathbf{G}}_i)^{-1}\mathbf{1}$ is unbounded as $\mathbf{\Gamma}$ tends to $\mathbf{I}$. Finally by the following lemma we can transfer this result to $\mathbf{G}_i - \overline{\mathbf{G}}_i$ yielding $\mathbf{G}_i \approx \overline{\mathbf{G}}_i$ whenever $\mathbf{\Gamma} \approx \mathbf{I}$:

Lemma 9.30: *Let $\mathbf{B} \in \mathbb{R}^{N\times N}$ be a non-negative invertible matrix. Suppose that there exists $c, d \in \mathbb{R}$ s.t.:*

$$c \leqslant \mathbf{B}^{-1}\mathbf{1} \leqslant d,$$

then:

$$\frac{1}{d} \leqslant \mathbf{B1} \leqslant \frac{1}{c} \tag{9.39}$$

Proof: Since $\mathbf{B}\mathbf{B}^{-1} = \mathbf{I}$, we have by multiplying both sides with $\mathbf{1}$:

$$\sum_j [\mathbf{B}]_{ij} \sum_k [\mathbf{B}^{-1}]_{jk} = 1, \quad \forall i \in [N].$$

Consequently we have for all $i \in [N]$:

$$1 = \sum_j [\mathbf{B}]_{ij}[\mathbf{B}^{-1}\mathbf{1}]_j \geqslant c\sum_j [\mathbf{B}]_{ij} = c[\mathbf{B1}]_i,$$

which gives the r.h.s. of the inequality (9.39). Now, by the similar computation we have the l.h.s. of the inequality (9.39):

$$1 = \sum_j [\mathbf{B}]_{ij}[\mathbf{B}^{-1}\mathbf{1}]_j \leqslant d \sum_j [\mathbf{B}]_{ij} = d[\mathbf{B1}]_i, \qquad \blacksquare$$

Specifically, if we have functions $\hat{\mathrm{f}}$ and $\check{\mathrm{f}}$ depending on $\mathbf{\Gamma}$, being unbounded as $\mathbf{\Gamma} \approx \mathbf{I}$, and satisfying:

$$\check{\mathrm{f}} \leqslant (\mathbf{G}_i - \overline{\mathbf{G}}_i)^{-1}\mathbf{1} \leqslant \hat{\mathrm{f}},$$

then it follows that:

$$\overline{\mathbf{G}}_i\mathbf{1} + \frac{1}{\hat{\mathrm{f}}} \leqslant \mathbf{G}_i\mathbf{1} \leqslant \overline{\mathbf{G}}_i\mathbf{1} + \frac{1}{\check{\mathrm{f}}},$$

and consequently:

$$\overline{\mathbf{h}}_i - \frac{1}{\check{\mathrm{f}}} \leqslant \mathbf{G}_i\mathbf{1} \leqslant \overline{\mathbf{h}}_i - \frac{1}{\hat{\mathrm{f}}},$$

where:

$$\overline{\mathbf{h}}_i = \left(1 - \frac{\eta^{\mathrm{T}}\mathbf{1}}{\mathrm{c}_i}\right)\mathbf{1} = \left(1 - \frac{\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1}}{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}}\right)\mathbf{1} = \frac{\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1}}{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}}\mathbf{1}.$$

Our choice for η_i and c_i is the following:

$$\eta_i^{\mathrm{T}} := \mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i) \quad \text{and} \quad \mathrm{c}_i := \mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1} \tag{9.40}$$

where $\mathbf{v}$ is the Perron-Frobenius left eigenvector of $\mathbf{A}$, i.e. the unique vector satisfying $\mathbf{v}^{\mathrm{T}}\mathbf{A} = \mathbf{v}^{\mathrm{T}}$. By this choice, it yields the following:

Lemma 9.31: *1.* $\mathbf{G}_i^{-1}\mathbf{1} = \mathbf{1} + (\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1}$

2. $1 \leqslant \mathbf{G}_i^{-1}\mathbf{1} \leqslant 1 + \frac{1-\gamma_{\min}^{(\backslash i)}}{1-\gamma_{\max}^{(i)}}\|\mathbf{A}_{i:}(\mathbf{I}-\mathbf{A}_{\backslash i})^{-1}\|_\infty$

Proof: For the first statement, notice that:

$$\begin{aligned}(\mathbf{I}-\tilde{\mathbf{C}}_i)\mathbf{1} &= (\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii} - \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{:i})\mathbf{1} \\ &= (\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{:i})\mathbf{1} \\ &= (\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1} + \mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1},\end{aligned}$$

where the last inequality follows from Lemma 9.13. Finally, by multiplying above expression with $(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}$, we obtain the first statement.

Now, for the second statement, notice that $(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})$ is non-negative. Consequently we have the l.h.s. inequality in the second statement. For the r.h.s. inequality in the second stament, we compute:

$$\begin{aligned}(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1} &\leqslant (1-\gamma_{\min}^{(\backslash i)})(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\mathbf{1}\\ &\leqslant (1-\gamma_{\min}^{(\backslash i)})\|\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{1}\\ &\leqslant \frac{1-\gamma_{\min}^{(\backslash i)}}{1-\gamma_{\max}^{(i)}}\|\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty\mathbf{1}.\end{aligned}$$

Furthermore, it is easy to see that:

$$\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1} \leqslant \mathbf{A}_{i:}(\mathbf{I}-\mathbf{A}_{\backslash i})^{-1}\mathbf{1}.$$

Combining both previous computations, we obtain the desired inequality. ∎

With the identity in above lemma, we can write d_i given in (9.38) as follows:

$$\begin{aligned}\mathrm{d}_i &= \mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\left[\mathbf{1}+(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1}\right]\\ &= \mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1}+\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})\mathbf{1}.\end{aligned} \tag{9.41}$$

The following identity of the second summand in above expression is of use for further step:

Lemma 9.32: *Let* $\mathbf{v}$ *be the Perron-Frobenius left eigenvector of* $\mathbf{A}$*. Then it holds:*

$$\mathbf{v}^{(i),\mathrm{T}}\mathbf{A}_{i:}[\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i}]^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\backslash i}] = \mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{\Gamma}_{\backslash i}^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\backslash i}]-\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})$$

Proof: Let be $\eta := \mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{v}$. We can write:

$$\eta_{i\leftrightarrow 1} := (\mathbf{Q}_{i\leftrightarrow 1}\eta)^{\mathrm{T}} = [\eta_i^{\mathrm{T}}, \eta_{\backslash i}^{\mathrm{T}}],$$

where $\eta_{\backslash i}^{\mathrm{T}} = \mathbf{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{v}^{(\backslash i)}$, and where $\mathbf{v}^{(\backslash i)}$ is the vector which is the result of deleting $\mathbf{v}^{(i)}$ entries from $\mathbf{v}$. Furthermore, for ease of notations, we write $\mathbf{B} := (\mathbf{I}-\mathbf{\Gamma})^{-1}(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})$ and:

$$\mathbf{B}_{i\leftrightarrow 1} := \mathbf{Q}_{i\leftrightarrow 1}(\mathbf{I}-\mathbf{\Gamma})^{-1}(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})\mathbf{Q}_{i\leftrightarrow 1} = \begin{bmatrix}(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}(\mathbf{I}-\mathbf{\Gamma}_i\mathbf{A}_{ii}) & -(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:}\\ -(\mathbf{I}-\mathbf{\Gamma}_{\backslash i})^{-1}\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{:i} & (\mathbf{I}-\mathbf{\Gamma}_{\backslash i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})\end{bmatrix}.$$

Now, we have:

$$\eta^{\mathrm{T}}\mathbf{B} = \mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma}\mathbf{A}) = \mathbf{v}^{\mathrm{T}}(\mathbf{\Gamma}^{-1}-\mathbf{A}) = \mathbf{v}^{\mathrm{T}}(\mathbf{\Gamma}^{-1}-\mathbf{I}) = \mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma}) = \eta^{\mathrm{T}},$$

that is η is the Perron eigenvector of the stochastic matrix $\mathbf{B}$. From this fact and the fact that $\mathbf{Q}_{i\leftrightarrow 1}\mathbf{Q}_{i\leftrightarrow 1} = \mathbf{I}$ and $\mathbf{Q}_{i\leftrightarrow 1}^{\mathrm{T}} = \mathbf{Q}_{i\leftrightarrow 1}$, we have:

$$\eta_{i\leftrightarrow 1}^{\mathrm{T}}\mathbf{B}_{i\leftrightarrow 1} = \eta^{\mathrm{T}}\mathbf{Q}_{i\leftrightarrow 1}\mathbf{Q}_{i\leftrightarrow}(\mathbf{I}-\mathbf{\Gamma})^{-1}(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})\mathbf{Q}_{i\leftrightarrow 1} = \eta^{\mathrm{T}}(\mathbf{I}-\mathbf{\Gamma})^{-1}(\mathbf{I}-\mathbf{\Gamma}\mathbf{A})\mathbf{Q}_{i\leftrightarrow 1} = \eta^{\mathrm{T}}\mathbf{Q}_{i\leftrightarrow 1} = \eta_{i\leftrightarrow 1}^{\mathrm{T}}.$$

Above identity asserts that:

$$\begin{aligned}
&-\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}\mathbf{A}_{i:} + \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})\\
&= -\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:} + \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})\\
&= -\eta_i^{\mathrm{T}}(\mathbf{I}-\mathbf{\Gamma}_i)^{-1}\mathbf{\Gamma}_i\mathbf{A}_{i:} + \eta_{\setminus i}^{\mathrm{T}}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})\\
&= \eta_{\setminus i}^{\mathrm{T}} = \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}).
\end{aligned} \tag{9.42}$$

Consequently:

$$\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}\mathbf{A}_{i:} = \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i}) - \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}).$$

By multiplying both sides to the right by $(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})$, we obtain the desired statement. ∎

In the following, we summarize our findings:

Lemma 9.33: *Let* $\mathbf{v}$ *be the Perron-Frobenius eigenvector of* $\mathbf{A}$. *Furthermore, Let* $\mathbf{G}_i$ *be given as in* (9.35) *and* $\overline{\mathbf{G}}_i$ *be a rank one matrix given as:*

$$\overline{\mathbf{G}}_i := \frac{\mathbf{1}\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)}{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})}.$$

Then we have:

$$(\mathbf{G}_i - \overline{\mathbf{G}}_i)^{-1}\mathbf{1} = \frac{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}}{\mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{1}}\mathbf{G}_i^{-1}\mathbf{1},$$

Proof: (9.37) asserts that:

$$(\mathbf{G}_i - \overline{\mathbf{G}}_i)^{-1}\mathbf{1} = \mathbf{G}_i^{-1}\mathbf{1}\left(1 + \frac{\frac{\mathrm{d}_i}{\mathrm{c}_i}}{1-\frac{\mathrm{d}_i}{\mathrm{c}_i}}\right) = \mathbf{G}_i^{-1}\mathbf{1}\frac{1}{1-\frac{\mathrm{d}_i}{\mathrm{c}_i}}, \tag{9.43}$$

where c_i is given as in (9.40) and d_i is given as in (9.38) with η_i is given as in (9.40). From (9.41) and Lemma 9.32, we obtain:

$$\begin{aligned}
\frac{\mathrm{d}_i}{\mathrm{c}_i} &= \frac{\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1} + \mathbf{v}^{(i),\mathrm{T}}\mathbf{A}_{i:}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})\mathbf{1}}{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}}\\
&= \frac{\mathbf{v}^{(i),\mathrm{T}}\mathbf{\Gamma}_i^{-1}(\mathbf{I}-\mathbf{\Gamma}_i)\mathbf{1} + \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}[\mathbf{I}-\mathbf{\Gamma}_{\setminus i}] - \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})}{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}}\\
&= \frac{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1} - \mathbf{v}^{(\setminus i),\mathrm{T}}\mathbf{\Gamma}_{\setminus i}^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})(\mathbf{I}-\mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}(\mathbf{I}-\mathbf{\Gamma}_{\setminus i})}{\mathbf{v}^{\mathrm{T}}\mathbf{\Gamma}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{1}},
\end{aligned}$$

and thus:

$$\frac{1}{1-\frac{\mathrm{d}_i}{\mathrm{c}_i}} = \frac{\mathbf{v}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\mathbf{I}-\boldsymbol{\Gamma})\mathbf{1}}{\mathbf{v}^{(\backslash i),\mathrm{T}}\boldsymbol{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})}.$$

Setting this into (9.43), we obtain the desired identity. ■

We can further estimate the r.h.s. of the identity in above Lemma and provide the following inequalities:

Lemma 9.34: *Consider the setting given in Lemma 9.33.*

$$\mathbf{G}_i\mathbf{1} \leqslant \overline{\mathbf{G}}_i\mathbf{1} + \frac{\gamma_{\max}^{(i)}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty}{\gamma_{\min}^{(\backslash i)}}\frac{(1-\gamma_{\min}^{(\backslash i)})^2}{1-\gamma_{\max}^{(i)}} \tag{9.44}$$

$$\mathbf{G}_i\mathbf{1} \geqslant \overline{\mathbf{G}}_i\mathbf{1} + \frac{\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_{\min}}{\gamma_{\max}^{(\backslash i)}\left(1+\frac{1-\gamma_{\min}^{(i)}}{1-\gamma_{\max}^{(i)}}\|\mathbf{A}_{i:}(\mathbf{I}-\mathbf{A}_{\backslash i})^{-1}\|_\infty\right)}\frac{(1-\gamma_{\max}^{(\backslash i)})^2}{\left(\frac{1-\gamma_{\min}^{(i)}}{\gamma_{\min}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}+\frac{1-\gamma_{\min}^{(\backslash i)}}{\gamma_{\min}^{(\backslash i)}}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}\right)} \tag{9.45}$$

Proof: We first show the first inequality. Notice that:

$$\begin{aligned}\mathbf{v}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\mathbf{I}-\boldsymbol{\Gamma})\mathbf{1} &= \mathbf{v}^{(i),\mathrm{T}}\boldsymbol{\Gamma}_i^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_i)\mathbf{1} + \mathbf{v}^{(\backslash i),\mathrm{T}}\boldsymbol{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})\mathbf{1} \geqslant \mathbf{v}^{(i),\mathrm{T}}\boldsymbol{\Gamma}_i^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_i)\mathbf{1}\\ &\geqslant \frac{1-\gamma_{\max}^{(i)}}{\gamma_{\max}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1},\end{aligned}$$

and that:

$$\mathbf{v}^{(\backslash i),\mathrm{T}}\boldsymbol{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})\mathbf{1} \leqslant \frac{(1-\gamma_{\min}^{(\backslash i)})^2}{\gamma_{\min}^{(\backslash i)}}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}.$$

Setting these estimations into the identity in Lemma 9.33, we obtain:

$$\begin{aligned}(\mathbf{G}_i-\overline{\mathbf{G}}_i)^{-1}\mathbf{1} &\geqslant \frac{\gamma_{\min}^{(\backslash i)}}{\gamma_{\max}^{(i)}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty}\frac{\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}}{\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}}\frac{1-\gamma_{\max}^{(i)}}{(1-\gamma_{\max}^{(\backslash i)})^2}\mathbf{G}_i^{-1}\mathbf{1}\\ &\geqslant \frac{\gamma_{\min}^{(\backslash i)}}{\gamma_{\max}^{(i)}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty}\frac{\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}}{\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}}\frac{1-\gamma_{\max}^{(i)}}{(1-\gamma_{\max}^{(\backslash i)})^2},\end{aligned}$$

where the second inequality follows from Lemma 9.31. From here, we have by Lemma 9.30:

$$(\mathbf{G}_i-\overline{\mathbf{G}}_i)\mathbf{1} \leqslant \frac{\gamma_{\max}^{(i)}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty}{\gamma_{\min}^{(\backslash i)}}\frac{\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}}{\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}}\frac{(1-\gamma_{\min}^{(\backslash i)})^2}{1-\gamma_{\max}^{(i)}},$$

and thus (9.44) is shown.

Now, we show (9.45). For this sake, we compute:

$$\begin{aligned}\mathbf{v}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\mathbf{I}-\boldsymbol{\Gamma})\mathbf{1} &= \mathbf{v}^{(i),\mathrm{T}}\boldsymbol{\Gamma}_i^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_i)\mathbf{1} + \mathbf{v}^{(\backslash i),\mathrm{T}}\boldsymbol{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})\mathbf{1}\\ &\leqslant \frac{1-\gamma_{\min}^{(i)}}{\gamma_{\min}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1} + \frac{1-\gamma_{\min}^{(\backslash i)}}{\gamma_{\min}^{(\backslash i)}}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1},\end{aligned}$$

and:

$$\mathbf{v}^{(\backslash i),\mathrm{T}}\boldsymbol{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})\mathbf{1} \geqslant \frac{(1-\gamma_{\max}^{(\backslash i)})^2}{\gamma_{\max}^{(\backslash i)}}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_{\min}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}$$

Setting these estimations into the identity in Lemma 9.33, we obtain:

$$\begin{aligned}(\mathbf{G}_i-\overline{\mathbf{G}}_i)^{-1}\mathbf{1} &\leqslant \frac{\gamma_{\max}^{(\backslash i)}}{\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_{\min}}\frac{\left(\frac{1-\gamma_{\min}^{(i)}}{\gamma_{\min}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}+\frac{1-\gamma_{\min}^{(\backslash i)}}{\gamma_{\min}^{(\backslash i)}}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}\right)}{(1-\gamma_{\max}^{(\backslash i)})^2}\mathbf{G}_i^{-1}\mathbf{1}\\ &\leqslant \frac{\gamma_{\max}^{(\backslash i)}\left(1+\frac{1-\gamma_{\min}^{(\backslash i)}}{1-\gamma_{\max}^{(i)}}\|\mathbf{A}_{i:}(\mathbf{I}-\mathbf{A}_{\backslash i})^{-1}\|_\infty\right)}{\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_{\min}}\frac{\left(\frac{1-\gamma_{\min}^{(i)}}{\gamma_{\min}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}+\frac{1-\gamma_{\min}^{(\backslash i)}}{\gamma_{\min}^{(\backslash i)}}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}\right)}{(1-\gamma_{\max}^{(\backslash i)})^2},\end{aligned}$$

where the second inequality follows from Lemma 9.31. Consequently, we have by Lemma 9.30:

$$(\mathbf{G}_i-\overline{\mathbf{G}}_i)\mathbf{1} \geqslant \frac{\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_{\min}}{\gamma_{\max}^{(\backslash i)}\left(1+\frac{1-\gamma_{\min}^{(\backslash i)}}{1-\gamma_{\max}^{(i)}}\|\mathbf{A}_{i:}(\mathbf{I}-\mathbf{A}_{\backslash i})^{-1}\|_\infty\right)}\frac{(1-\gamma_{\max}^{(\backslash i)})^2}{\left(\frac{1-\gamma_{\min}^{(i)}}{\gamma_{\min}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}+\frac{1-\gamma_{\min}^{(\backslash i)}}{\gamma_{\min}^{(\backslash i)}}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}\right)} \qquad \blacksquare$$

Finally, by combining the relation (9.36) and above Lemma, we obtain the following approximation statement for the measure $\mathbf{h}_i$ of our interest:

Theorem 9.35: *It holds:*

$$\overline{\mathbf{h}}_i - \check{\mathrm{f}}(\boldsymbol{\Gamma}) \leqslant \mathbf{h}_i \leqslant \overline{\mathbf{h}}_i - \hat{\mathrm{f}}(\boldsymbol{\Gamma}),$$

where:

$$\overline{\mathbf{h}}_i := \frac{\mathbf{v}^{(\backslash i),\mathrm{T}}\boldsymbol{\Gamma}_{\backslash i}^{-1}(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i})\mathbf{1}}{\mathbf{v}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\mathbf{I}-\boldsymbol{\Gamma})\mathbf{1}}\mathbf{1}, \tag{9.46}$$

with $\mathbf{v}$ denoting the Perron-Frobenius eigenvector of $\mathbf{A}$, and where:

$$\begin{aligned}\hat{\mathrm{f}}(\boldsymbol{\Gamma}) &= \frac{\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_{\min}}{\gamma_{\max}^{(\backslash i)}\left(1+\frac{1-\gamma_{\min}^{(i)}}{1-\gamma_{\max}^{(i)}}\|\mathbf{A}_{i:}(\mathbf{I}-\mathbf{A}_{\backslash i})^{-1}\|_\infty\right)}\frac{(1-\gamma_{\max}^{(\backslash i)})^2}{\left(\frac{1-\gamma_{\min}^{(i)}}{\gamma_{\min}^{(i)}}\mathbf{v}^{(i),\mathrm{T}}\mathbf{1}+\frac{1-\gamma_{\min}^{(\backslash i)}}{\gamma_{\min}^{(\backslash i)}}\mathbf{v}^{(\backslash i),\mathrm{T}}\mathbf{1}\right)}\\ \check{\mathrm{f}}(\boldsymbol{\Gamma}) &= \frac{\gamma_{\max}^{(i)}\|(\mathbf{I}-\boldsymbol{\Gamma}_{\backslash i}\mathbf{A}_{\backslash i})^{-1}\|_\infty}{\gamma_{\min}^{(\backslash i)}}\frac{(1-\gamma_{\min}^{(\backslash i)})^2}{1-\gamma_{\max}^{(i)}}.\end{aligned}$$

Now, we close the chapter by further discussing Theorem (9.35). Our interest is to see for which susceptibility constants near 1 (or equivalently bias constant near 0) $\mathbf{h}_i$ can be approximated by $\overline{\mathbf{h}}_i$ given in (9.46). For this sake, we estimate the functions $\hat{\mathrm{f}}$ and $\check{\mathrm{f}}$ given in Theorem (9.35). We have obviously:

$$\|(\mathbf{I} - \mathbf{\Gamma}_{\setminus i}\mathbf{A}_{\setminus i})^{-1}\|_{\min} \geqslant 1.$$

Consequently, if:

$$\frac{1 - \gamma_{\min}^{(i)}}{1 - \gamma_{\max}^{(i)}} = \Omega(1),$$

we have that:

$$\hat{\mathrm{f}}(\mathbf{\Gamma}) \geqslant C_1 \frac{(1 - \gamma_{\max}^{(\setminus i)})^2}{1 - \gamma_{\min}^{(i)}},$$

for a constant $C_1 > 0$. Furthermore, it clearly holds:

$$\check{\mathrm{f}}(\mathbf{\Gamma}) \leqslant C_2 \frac{(1 - \gamma_{\min}^{(\setminus i)})^2}{1 - \gamma_{\max}^{(i)}},$$

for a constant $C_2 > 0$. So from previous observations, we have for instance that if the susceptibility constants of the agents in cluster i is of order $\Theta(1 - \epsilon^p)$, where $p > 0$, and if the susceptibility constants of the agents out of cluster i is of order $\Theta(1 - \epsilon^q)$, where $q > p/2$, then $\mathbf{h}_i \approx \overline{\mathbf{h}}_i$ with small error of order $\Theta(\epsilon^{2q-p})$.

10. Summary, Conclusions, and Outlook

10.1. Part I: Resource Sustainable Robust Online Learning in Games

In the first part of this thesis, we considered systems of selfish learning agents. We used the online learning paradigm and game theory to model their strategic interaction. Main aspects of our interest in the first part of this thesis were:

- the effect of the lack of agents' global view to the population's outcome,
- the effect of uncertainty in the information/feedback obtained by the agents to the system,
- and the sustainability of the population's behaviour in face of selfishness of the agents.

We were able to provide not only results of asymptotic nature, which is usual in the literature of game theory and dynamical systems, but also results of non-asymptotic nature. In the following we provide a more comprehensive summaries and conclusions of the investigations made in the first part of the thesis. Also, we present at each of the Subsections some interesting research directions for future works.

10.1.1. Chapter 3: On the Convergence of Online Mirror Descent for Aggregative Games with Approximated Aggregates

In Chapter 3, we consider the setting of the aggregative games, which occurs in a vast number of applications such as signal processing and cummunications, smart grid, competitive markets and congestion control for networks. Assuming that the agents learn in the online manner via first-order feedback, we propose a sufficient condition on the degree of the uncertainties of the agents about the aggregate for the convergence of the population's dynamic to the Wardrop equilibrium. Not only for analyzing the dynamic occuring in practical applications proposed in this work, our results can also of course be used to find a Wardrop equilibrium of a given game.

In order to illustrate the possible decay behaviour of the error between the aggregate and its estimates implied by our main result, one may consider the case where:

$$\gamma_k = \mathcal{O}\left(\frac{1}{k^\alpha}\right), \quad \text{with} \quad \alpha > \frac{1}{2}.$$

If:

$$\|\sigma_k^{(i)} - \sigma(\boldsymbol{X}_k)\|_{\tilde{V}} = \mathcal{O}\left(\frac{1}{k^\beta}\right), \quad \text{where} \quad \beta > -\alpha + 1,$$

then (3.12) is fulfilled. Moreover, by choosing the step size:

$$\gamma_k = \mathcal{O}\left(\frac{1}{k\log(k)}\right),$$

it is even possible to handle the slow decay:

$$\|\sigma_k^{(i)} - \sigma(\boldsymbol{X}_k)\|_{\tilde{V}} = \mathcal{O}\left(\frac{1}{\log(k)}\right).$$

Those observations show in particular the role of the step size choice in handling the negative effect of the agents' uncertainties about the value of the aggregate, to the convergence behaviour.

Notice that the condition (3.12) leading to the convergence of our proposed algorithm is a sufficient condition. So one may ask whether the given condition is necessary. What we can say in the direction of the answer of this question, is that the condition (3.12) is reasonable, since a necessary requirement for (3.12) to hold is that the aggregate and its estimates coincides in the time limit. However, we think that one can replace (3.12) by the weaker condition:

$$\lim_{n\to\infty} \frac{\sum_{k=0}^n \gamma_k \|\hat{\sigma}_k^{(i)} - \sigma(\boldsymbol{X}_k)\|_{\tilde{V}} < \infty}{\sum_{k=0}^n \gamma_k} = 0.$$

One assumption we made in this work is that the step size sequences of the agents coincide which is quiet restrictive. So in the future work, we aim to handle the case where the step size sequences of the agents differ.

Since we basically work with the gradient operator of the utility function and the corresponding variational inequality, attentive readers would notice that the notion of concavity can be replaced by the notion of pseudo-concavity. Moreover, if the notion of monotonicity is replaced by the notion of strong monotonicity, we believe that the the convergence rate can be specified. This line of work will also be exploited in the future.

If the aggregate is the mean of agents' actions, one can by means of our technique and the aggregate estimation technique given in [77] give a distributed online mirror ascent algorithm to achieve a Wardrop equilibrium. Beside providing a more detailed analysis of this aspect, we plan in the future work to handle aggregates other than the mean aggregate.

10.1.2. Chapter 4: Coordinated Online Learning for Multi-Agent Systems with Coupled Constraints and Perturbed Utility Observations

In Chapter 4, we considered the setting of general games with the feature of constraints coupled among the agents. This sort of constraints arise in different practical applications such as those where the agents are competing on the utilization of limited resources. We have proposed a novel decentralized pricing method that aims to encourage resource sustainable behavior in a population of selfish online learning agents having noisy first-order feedback by leading them toward a generalized Nash equilibrium of the game with corresponding coupled constraints.

The given results are based on the assumption that the utility functions and the constraint violation functions are continuously differentiable and that the utility function is strictly convex. However, it is straightforward to generalize the results in order to handle (not-necessarily continuously differentiable) convex function by replacing the gradients with subgradients, and by involving, besides, the notion of convergence to a set (see [76]). The latter is necessary since the set of variational Nash equilibrium $\mathrm{SOL}(\mathcal{Q}, \mathbf{v})$ in the convex utility case is in general not a singleton. The simplification made in this work is only for the sake of readability.

In the case that the martingale noise is persistent, the method gives a.s. convergence of the population's state to the generalized Nash equilibrium and consequently a.s. compliance of the resource constraints in the asymptotic limit for the rich class of polynomially decaying step-size policies of order $\gamma_n = \Theta(n^{-b})$ where $b \in (0, 1]$. However, we were only able to give this guarantee for the ergodic average of the population iterates. For the indeed population's iterate, we only could show the a.s. convergence for $b \in (1/2, 1]$. Thus the case $b \in (0, 1/2]$ remains open.

In case where $\gamma_n = \Theta(n^{-1/2})$, we were able to provide in the persistent noise case a non-asymptotic decaying expectation bound of order $\mathcal{O}(\ln^{3/2}(n)/\sqrt{n})$ for the amount of resource constraints violations caused by the ergodic average of the population's iterates. This decay rate matches, up to the ln-factor, with the fundamental limit of the convergence speed described by the lower complexity bound for black-box subgradient methods (see Theorem 3.2.1 in [204]). Thus we expect that our result is optimal.

Another interesting occurrence we observed, both in the theoretical and numerical investigations, is that the choice of mirror map might have an impact on the dimensional dependence of the quality of MAARP. This effect is rarely considered in the literature of the Nash equilibrium finding since it mostly uses the Euclidean projection as the mapping which realizes the first-order update in the feasible strategy set (to name a few: [84, 104, 205]). In future work, we aim to exploit this aspect further.

We also provided a non-asymptotic expectation - and high probability bound for the distance between the ergodic average of the primal-dual equilibrium gap of the noisy MDAL and the variationally nash equilibrium. An implication of this result is that if the noise is persistent with "power" σ^2, a bound of order $\mathcal{O}(\sigma/\sqrt{n})$ for fixed time horizon and constant step size and a bound of order $\mathcal{O}(\sigma^2(\ln(n)/\sqrt{n}))$ for variable stepsize is achievable.

10.1.3. Chapter 5: Impact of Agents' Price Sensitivity on the Resource Sustainable Pricing

In the Chapter 5 we extended the algorithm given in Chapter 4 by introducing an additional parameter β specifying extrinsically the price sensitivity of the agents. The discussions tantalized by this procedure provides an interesting insight into the answer of the question of how much amount of control do a non-cooperative system needed in order that a specific population's goal is achieved. Our overall answer is that, to a certain degree, sensitizing the agents regarding to the control variable, can encourage the selfish agents working toward a mutual goal. However, sacrificing agents selfishness by a high degree of control can be in contrary to the desired global goal.

If $\beta = 1$, then the algorithm introduced in Chapter 5 is basically the MAARP given in Chapter 4. We have shown that $\beta = 2$, meaning that the weight of the price in agents' decision strategy is two times usual one, then the corresponding pricing mechanism foster the resource sustainable behaviour. This is shown by the non-asymptotic bound we provide in Chapter 5. Although the bound is comparable to the given in Chapter 4 respective to the order, it is more convenient than the bound given in Chapter 4.

By numerical experiments, we are able to show that the choice $\beta = 2$ is superior to the pricing mechanism given in Chapter 4, indication that the bound given in this chapter can be improved. In this work we plan to exploit this aspect theoretically. Also we have shown by simulations that if the sensitivity parameter is too high, i.e. higher than 2, then the performance of the corresponding mechanism become worst, giving the hint that excessive control by prices can be in opposite to the desired goal. It remains to show the latter theoretically. We plan to do the latter in the future work.

10.1.4. Chapter 6: Resource-Aware Control via Pricing for Congestion Game with Finite-Time Guarantees

Assuming that the agents are choosing their action based on the average historical cost of the resource bundles and the logit choice rule, we introduce in this work a resource-centric pricing mechanism which allows a non-asymptotic guarantee of the sub-linear growth of the expected aggregated violation of the resource constraints of order $\mathcal{O}(\sqrt{n})$. In case that the resource contraints are not overly strict, we observe numerically that the resource sustainability delivered by our method, does not come with significant discrimination of the agents. For the general case, trade-off effect between resource sustainability and population's welfare might occur. In the future, we plan to explain this aspects formally.

10.2. Part III: Distributed Coordination Algorithms

In the last part of this thesis, we concerned again with the setting of multiple agents. In contrast to the first part of this thesis, we assume that the agents has the ability to locally communicate. Main aspects of our interest:

- the efficiency of the continuous-time distributed gradient method in face of gradient noise,
- and the impact of informational bias to the opinion formation.

In the following we provide a summary and conclusion of the results of our findings:

10.2.1. Chapter 8: Stochastic Dynamic of First-Order Flocking-based Distributed Optimization

In Chapter 8, we have analyzed a continuous-time stochastic distributed gradient descent implementation which based on flocking with pure repulsion type and whose gradient noise is Gaussian.

We have estimated the error between the corresponding proposal solution of each agents and the indeed optimizer both in expectation and also in probability. The corresponding bounds reflects both roles of the step size/gradient weight α acting antagonistically, as the parameter dictating the convergence rate and as the parameter reducing the effect of the accumulated deviation caused by noise (cf. (8.21)). We notice that the communication strength β and the network structure $\lambda_2(\mathbf{L})$ reduces the negative influence of dimension and number of agents to the cohesiveness of the iterate of each agents which has in turn positive influence to the correctness of the proposal solution of each agents. Another factor which robustifies the algorithm is the strong convexity of the function, which together with the step size is able to annihilate the influence of instantaneous noise (cf. condition (8.22)). We have discussed the trade-off between the parameters by considering cases where the noise is persistent and/or vanishing.

We believe that under certain condition on the step size and noise volatility, the condition (8.20) of boundedness of the gradient on the average path of the solution of (8.9) is unnecessary to state. We aim to investigate this aspect in the future work. Another direction for the future work is to extend the algorithm and analysis to the mirror descent and Nesterov's acceleration case.

10.2.2. Chapter 9: Mesoscopic Stability of the Friedkin-Johnsen Opinion Dynamics

In Chapter 9, we have provided an analysis of mixing property of the opinion dynamic of a population of agents in face of disturbance by informational bias. We have observed that the latter factor can cause the clusterness of the opinion of the population of agents.

We have introduced the notion of substochastic complementation giving an efficient approximation of the opinion dynamic by cluster-independent dynamic. Based on this notion, we have provided a novel measure for the degree of the cluster-dispersion of the opinions caused by the informational bias. Also, we have analyzed the opinion behavior for several limit cases of informational bias influence. Our conclusions is as follows:

- In case that the opinion of in-cluster agents almost determined by informational bias, the dynamic of the in-cluster agents is almost uncoupled of the other agents irrespective of the out-cluster agents' behavior. In contrast, if the out-cluster agents are under a great influence of informational bias, the degree of the uncoupling of the cluster dynamic depends on the own structure of the cluster.
- In case that the influence of informational bias is negligible, the degree of clusterness of the opinion dynamic can be arbitrarily. However, it can be approximated by the quantity depending on the susceptibility constant and the eigenvector corresponds to the eigenvalue with the largest absolute value.

Notice that one can also make advantage from this approximation by using it instead of the whole network for further tasks such as e.g. further analysis network statistics, since working with whole network is in some cases, e.g., in the case that the network is of large-scale, complicated. Also this kind of approximation gives a more detailed and concise view of the network. For instance, it allows one to find agents, which are (Katz-Bonacich) central in the corresponding cluster, instead of only one agent central in the whole network [206, 207]. In the future work, we aim to exploit this aspects more in detail.

Bibliography

[1] A. C. Pigou, *The Economics of Welfare*. London: Macmillan, 1924.

[2] M. O. Jackson and Y. Zenou, "Games on Networks," in *Handbook of Game Theory with Economic Applications*, pp. 95–163, Elsevier, 2015.

[3] B. Marin and R. Mayntz, eds., *Policy Networks: Empirical Evidence and Theoretical Considerations,*. Campus Verlag, 1991.

[4] A. Sandström, *Policy Networks: The relation between structure and performance.* PhD thesis, Luleå University of Techchnology, 2008.

[5] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292 – 2330, 2008.

[6] H. Wang, H. Zhao, J. Zhang, D. Ma, J. Li, and J. Wei, "Survey on Unmanned Aerial Vehicle Networks: A Cyber Physical System Perspective," *IEEE Communications Surveys Tutorials*, vol. 22, no. 2, pp. 1027–1070, 2020.

[7] T. Suzuki, T. Sekine, T. Fujii, H. Asama, and I. Endo, "Cooperative formation among multiple mobile robot teleoperation in inspection task," in *Proc. of the 39th IEEE CDC*, vol. 1, 2000.

[8] P. R. Wurman, R. D'Andrea, and M. Mountz, "Coordinating hundreds of cooperative, autonomous vehicles in warehouses," *AI Magazine*, vol. 29, no. 1, 2008.

[9] J. Alonso-Mora, E. Montijano, T. Nägeli, O. Hilliges, M. Schwager, and D. Rus, "Distributed multi-robot formation control in dynamic environments," *Autonomous Robot*, vol. 43, pp. 1079–1100, 2019.

[10] A. Sargolzaei, A. Abbaspour, and C. D. Crane, *Optimization, Learning, and Control for Interdependent Complex Networks*, ch. Control of Cooperative Unmanned Aerial Vehicles: Review of Applications, Challenges, and Algorithms, pp. 229–255. Cham: Springer International Publishing, 2020.

[11] M. Hashmi, S. Hänninen, and K. Mäki, "Survey of smart grid concepts, architectures, and technological demonstrations worldwide," in *2011 IEEE PES ISGT LA*, pp. 1–7, 2011.

[12] A. H. Al-Badi, R. Ahshan, N. Hosseinzadeh, R. Ghorbani, and E. Hossain, "Survey of smart grid concepts and technological demonstrations worldwide emphasizing on the oman perspective," *Applied System Innovation*, vol. 3, no. 1, p. 5, 2020.

[13] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press, 2012.

[14] O. Said and M. Masud, "Towards internet of things: Survey and future vision," *International Journal of Computer Networks*, vol. 5, no. 1, pp. 1–17, 2013.

[15] S. E. Bibri and J. Krogstie, "Smart sustainable cities of the future: An extensive interdisciplinary literature review," *Sustainable Cities and Society*, vol. 31, pp. 183 – 212, 2017.

[16] M. F. Othman and K. Shazali, "Wireless sensor network applications: A study in environment monitoring system," *Procedia Engineering*, vol. 41, pp. 1204 – 1210, 2012. International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012).

[17] J. Manathara, P. B. Sujit, and R. W. Beard, "Multiple UAV Coalitions for a Search and Prosecute Mission," *J. Intell. Robot Syst,*, vol. 62, pp. 125–158, 2011.

[18] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on mri," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102 – 127, 2019. Special Issue: Deep Learning in Medical Physics.

[19] J. Wiens and E. S. Shenoy, "MachineLearning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology," *Clin. infect. dis.*, vol. 66, no. 1, 2018.

[20] D. Davies, "How Machine Learning in Search Works: Everything You Need to Know." Search Engine Jour., 2020.

[21] N. Jones, "Computer science: The learning machines," *Nature News*, vol. 505, no. 7482, p. 146, 2014.

[22] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.

[23] P. Stone and M. Veloso, "Multiagent Systems: A Survey from a Machine Learning Perspective," *Aut. Rob.*, vol. 8, pp. 345–383, 2000.

[24] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Rob. and aut. sys.*, vol. 57, no. 5, pp. 469–483, 2009.

[25] M. Zinkevich, "Online Convex Programming and Generalized Infinitesimal Gradient Ascent," in *Proceedings of the 20th International Conference on International Conference on Machine Learning*, pp. 928 – 935, 2003.

[26] B. Narayanaswamy, V. K. Garg, and T. S. Jayram, "Online optimization for the smart (micro) grid," in *2012 Third International Conference on Future Systems: Where Energy, Computing and Communication Meet (e-Energy)*, pp. 1–10, 2012.

[27] M. Moeini-Aghtaie, P. Dehghanian, M. Fotuhi-Firuzabad, and A. Abbaspour, "Multiagent Genetic Algorithm: An Online Probabilistic View on Economic Dispatch of Energy Hubs Constrained by Wind Availability," *IEEE Transactions on Sustainable Energy*, vol. 5, pp. 699–708, Apr. 2014.

[28] M. Lin, A. Wierman, L. L. H. Andrew, and E. Thereska, "Dynamic right-sizing for power-proportional data centers," in *2011 Proceedings IEEE INFOCOM*, Apr. 2011.

[29] M. Lin, Z. Liu, A. Wierman, and L. L. H. Andrew, "Online algorithms for geographical load balancing," in *2012 International Green Computing Conference (IGCC)*, pp. 1–10, June 2012.

[30] T. Chen, Q. Ling, and G. B. Giannakis, "An Online Convex Optimization Approach to Proactive Network Resource Allocation," *IEEE Transactions on Signal Processing*, vol. 65, pp. 6350–6354, Dec. 2017.

[31] L. Gan, A. Wierman, U. Topcu, N. Chen, and S. H. Low, "Real-time Deferrable Load Control: Handling the Uncertainties of Renewable Generation," in *Proceedings of the Fourth International Conference on Future Energy Systems*, pp. 113–124, 2013.

[32] S. Kim and G. B. Giannakis, "Real-time electricity pricing for demand response using online convex optimization," pp. 1–5, Feb. 2014.

[33] V. Joseph and G. de Veciana, "Jointly optimizing multi-user rate adaptation for video transport over wireless systems: Mean-fairness-variability tradeoffs," in *2012 Proceedings IEEE INFOCOM*, Mar. 2012.

[34] F. Zanini, D. Atienza, G. De Micheli, and S. Boyd, "Online convex optimization-based algorithm for thermal management of mpsocs," in *Proc. of the 20th Great lakes symp. on VLSI*, pp. 203–208, ACM, 2010.

[35] T. Chen, Q. Ling, Y. Shen, and G. B. Giannakis, "Heterogeneous Online Learning for Thing-Adaptive Fog Computing in IoT," *IEEE Internet of Things Journal*, vol. 5, pp. 4328 – 4341, Dec. 2018.

[36] T. Chen and G. B. Giannakis, "Bandit Convex Optimization for Scalable and Dynamic IoT management," *IEEE Internet of Things Journal*, vol. 6, pp. 1276–1286, Feb. 2019.

[37] T. Chen, S. Barbarossa, X. Wang, G. Giannakis, and Z.-L. Zhang, "Learning and Management for Internet of Things: Accounting for Adaptivity and Scalability," *Proc. of the IEEE*, vol. 107, pp. 778–796, 2019.

[38] S. Shalev-Shwartz, "Online Learning and Online Convex Optimization," *Found. and Tr. in Mach. Lear.*, vol. 4, no. 2, pp. 107 – 194, 2012.

[39] S. Bubeck and N. Cesa-Bianchi, "Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems," *Found. and Tr. in Mach. Lear.*, vol. 5, no. 1, pp. 1–122, 2012.

[40] E. V. Belmega, P. Mertikopoulos, R. Negrel, and L. Sanguinetti, "Online convex optimization and no-regret learning: Algorithms, guarantees and applications," *arXiv:1804.04529*, 2018.

[41] G. Scutari, D. P. Palomar, F. Facchinei, and J. Pang, "Convex Optimization, Game Theory, and Variational Inequality Theory," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 35–49, 2010.

[42] P. Mertikopoulos and A. L. Moustakas, "Learning in an Uncertain World: MIMO Covariance Matrix Optimization With Imperfect Feedback," *IEEE Transactions on Signal Processing*, vol. 64, no. 1, pp. 5–18, 2016.

[43] J. Abernethy and A. Rakhlin, "Optimal strategies and minimax lower bounds for online convex games," in *Proc. of 19th COLT*, 2008.

[44] W. Wang and M. Á. Carreira-Perpiñán, "Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application," *CoRR*, vol. abs/1309.1541, 2013.

[45] S. M. Kakade, S. Shalev-Shwartz, and A. Tewari, "Regularization techniques for learning with matrices," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 1865–1890, 2012.

[46] P. Mertikopoulos, E. V. Belmega, R. Negrel, and L. Sanguinetti, "Distributed stochastic optimization via matrix exponential learning," *IEEE Transactions on Signal Processing*, vol. 65, pp. 2277–2290, May 2017.

[47] Y. Nesterov, "Primal-dual subgradient methods for convex problems," *Math. Prog.*, vol. 120, pp. 221–259, Aug 2009.

[48] A. Nemirovski, A. Juditsky, G. Lan, and A. . Shapiro, "Robust Stochastic Approximation Approach to Stochastic Programming," *SIAM J. on Opt.*, vol. 19, pp. 1574 – 1609, Jan. 2008.

[49] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, 1970.

[50] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, vol. 317 of *A Ser. of Comp. Stud. in Math.* Springer-Verlag, 1998.

[51] P. Mertikopoulos and W. H. Sandholm., "Learning in games via reinforcement and regularization," *Math. of Op. Res.*, vol. 14, no. 1, pp. 124 – 143, 2016.

[52] W. H. Sandholm, *Evolutionary Game Theory*, pp. 1000–1029. New York, NY: Springer New York, 2012.

[53] J. Newton, "Evolutionary game theory: A renaissance," *Games*, vol. 9, no. 2, 2018.

[54] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[55] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer-Verlag New York, 2003.

[56] P. Hall and C. Heyde, *Martingale Limit Theory and its Application.* Probability and Mathematical Statistics: A Series of Monographs and Textbooks, Academic Press, 1980.

[57] E. A. Jorswieck, E. G. Larsson, M. Luise, and H. V. Poor, "Game theory in signal processing and communications," *IEEE Sig. Proc. Mag.*, vol. 26, pp. 17–132, Sep. 2009.

[58] A. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid," *IEEE Trans. on Smart Grid*, vol. 1, pp. 320 – 331, Dec. 2010.

[59] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Sig. Proc. Mag.*, vol. 29, no. 5, pp. 86 – 105, 2012.

[60] R. Deng, Z. Yang, J. Chen, N. R. Asr, and M. Chow, "Residential Energy Consumption Scheduling: A Coupled-Constraint Game Approach," *IEEE Trans. on Smart Grid*, vol. 5, pp. 1340 – 1350, May 2014.

[61] S. Li, W. Zhang, J. Lian, and K. Kalsi, "Market-Based Coordination of Thermostatically Controlled Loads – part I: A Mechanism Design Formulation," *IEEE Trans. on Pow. Sys.*, vol. 31, pp. 1170 – 1178, March 2016.

[62] S. Grammatico, B. Gentile, F. Parise, and J. Lygeros, “A mean field control approach for demand side management of large populations of thermostatically controlled loads,” in *Proc. of the IEEE Eur. Cont. Conf.*, 2015.

[63] Z. Ma, D. S. Callaway, and I. A. Hiskens, “Decentralized charging control of large populations of plug-in electric vehicles,” *IEEE Trans. on Cont. Sys. Tech.*, vol. 21, pp. 67 – 78, Jan. 2013.

[64] F. Parise, M. Colombino, S. Grammatico, and J. Lygeros, “Mean field constrained charging policy for large populations of plug-in electric vehicles,” in *53rd IEEE Conf. on Dec. and Cont. (CDC)*, pp. 5101 – 5106, Dec. 2014.

[65] Z. Ma, S. Zou, L. Ran, X. Shi, and I. A. Hiskens, “Efficient decentralized coordination of large-scale plug-in electric vehicle charging,” *Aut.*, vol. 69, pp. 35 – 47, 2016.

[66] S. Grammatico, “Exponentially convergent decentralized charging control for large populations of plug-in electric vehicles,” in *IEEE 55th Conf. on Dec. and Cont. (CDC)*, pp. 5775 – 5780, Dec. 2016.

[67] N. Li, L. Chen, and M. A. Dahleh, “Demand Response Using Linear Supply Function Bidding,” *IEEE Trans. on Smart Grid*, vol. 6, pp. 1827–1838, July 2015.

[68] J. Barrera and A. Garcia, “Dynamic Incentives for Congestion Control,” *IEEE Trans. on Aut. Cont.*, vol. 60, pp. 299 – 310, Feb. 2015.

[69] N. S. Kukushkin, “Best response dynamics in finite games with additive aggregation,” *Games and Economic Behavior*, vol. 48, no. 1, pp. 94 – 110, 2004.

[70] M. Jensen, “Aggregative games and best-reply potentials,” *Economic Theory*, vol. 43, pp. 45–66, Apr 2010.

[71] P. Dubey, O. Haimanko, and A. Zapechelnyuk, “Strategic complements and substitutes, and potential games,” *Games and Economic Behavior*, vol. 54, no. 1, pp. 77 – 94, 2006.

[72] M. Smith, “The existence, uniqueness and stability of traffic equilibria,” *Transportation Research Part B: Methodological*, vol. 13, no. 4, pp. 295 – 304, 1979.

[73] T. Alpcan, T. Basar, R. Srikant, and E. Altman, “Cdma uplink power control as a noncooperative game,” in *40th IEEE Conf. on Dec. and Cont. (CDC)*, vol. 1, pp. 197 – 202, 2001.

[74] Z. Ma, D. S. Callaway, and I. A. Hiskens, “Decentralized charging control of large populations of plug-in electric vehicles,” *IEEE Trans. on Cont. Sys. Tech.*, vol. 21, pp. 67 – 78, Jan. 2013.

[75] R. Johari, S. Mannor, and J. N. Tsitsiklis, "Efficiency loss in a network resource allocation game: the case of elastic supply," *IEEE Trans. on Aut. Cont.*, vol. 50, pp. 1712–1724, Nov. 2005.

[76] P. Mertikopoulos and M. Staudigl, "On the convergence of gradient-like flows with noisy gradient input," *SIAM Journal on Optimization*, vol. 28, no. 1, pp. 163–197, 2018.

[77] J. Koshal, A. Nedić, and U. V. Shanbhag, "Distributed algorithms for aggregative game on graphs," *Op. Res.*, vol. 64, no. 3, pp. 680 – 704, 2016.

[78] A. Haurie and P. Marcotte, "On the relationship between nash—cournot and wardrop equilibria," *Networks*, vol. 15, no. 3, pp. 295–308, 1985.

[79] S. Dafermos and A. Nagurney, "Oligopolistic and competitive behavior of spatially separated markets," *Regional Science and Urban Economics*, vol. 17, no. 2, pp. 245 – 254, 1987.

[80] E. Altman and L. Wynter, "Equilibrium, games, and pricing in transportation and telecommunication networks," *Networks and Spatial Economics*, vol. 4, pp. 7–21, Mar 2004.

[81] E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter, "A survey on networking games in telecommunications," *Computers & Operations Research*, vol. 33, no. 2, pp. 286 – 311, 2006. Game Theory: Numerical Methods and Applications.

[82] E. Altman, R. Combes, Z. Altman, and S. Sorin, "Routing games in the many players regime," in *Proc. of the 5th Int. ICST Conference on Perf. Eval. Meth. and Tools*, pp. 525 – 527, 2011.

[83] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized Convergence to Nash Equilibria in Constrained Deterministic Mean Field Control," *IEEE Trans. on Aut. Cont.*, vol. 61, pp. 3315 – 3329, Nov. 2016.

[84] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J.Lygeros, "Nash and Wardrop Equilibria in Aggregative Games with Coupling Constraints," *IEEE Trans. on Aut. Cont.*, 2017. arXiv:1702.08789.

[85] X. Liu, R. Zhu, B. Jalaian, and Y. Sun, "Dynamic spectrum access algorithm based on game theory in cognitive radio networks," *Mob. Netw. Appl.*, vol. 20, pp. 817 – 827, Dec. 2015.

[86] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang, *Monotone Games for Cognitive Radio Systems*, pp. 83–112. London: Springer London, 2012.

[87] G. Abbas, Z. Halim, and Z. H. Abbas, "Fairness-driven queue management: A survey and taxonomy," *IEEE Communications Surveys Tutorials*, vol. 18, pp. 324 – 367, Firstquarter 2016.

[88] C. Borch, "High-frequency Trading, Algorithmic Finance, and the Flash Crash: Reflections on Eventalization," *Economy and Society*, vol. 45, pp. 350 – 378, Jan. 2016.

[89] D. Fudenberg, *The Theory of Learning in Games.* MIT Press, 1998.

[90] K. Zhang, Z. Yang, and T. Başar, "Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms," *arXiv:1911.10635*, 2019.

[91] B. Duvocelle, P. Mertikopoulos, M. Staudigl, and D. Vermeulen, "Learning in time-varying games," *CoRR*, vol. abs/1809.03066, 2018.

[92] Z. Zhou, P. Mertikopoulos, N. Bambos, P. Glynn, and C. Tomlin, "Multi-agent online learning with imperfect information," *Submitted.*, 2018.

[93] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *4OR*, vol. 5, pp. 173–210, Sep 2007.

[94] L. Pavel, "An Extension of Duality to a Game-theoretic Framework," *Automatica*, vol. 43, pp. 226–237, Feb. 2007.

[95] H. Yin, U. V. Shanbhag, and P. G. Mehta, "Nash Equilibrium Problems With Scaled Congestion Costs and Shared Constraints," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1702–1708, 2011.

[96] A. A. Kulkarni and U. V. Shanbhag, "On the variational equilibrium as a refinement of the generalized Nash equilibrium," *Automatica*, vol. 48, no. 1, pp. 45 – 55, 2012.

[97] G. Arslan, M. F. Demirkol, and S. Yüksel, "On Games with Coupled Constraints," in *51st IEEE CDC*, pp. 7151–7156, 2012.

[98] R. Johari, *Efficiency loss in market mechanisms for resource allocation.* PhD thesis, MIT, 2004.

[99] T. Roughgarden and É. Tardos, "Bounding the inefficiency of equilibria in nonatomic congestion games," *Games and Economic Behavior*, vol. 47, no. 2, pp. 389 –203, 2004.

[100] C. Papadimitriou, "Algorithms, games, and the internet," in *Proc. of 33th ACM on the Theory of Computing*, pp. 749 – 753, 2001.

[101] A. A. Kulkarni, "The Efficiency of Generalized Nash and Variational Equilibria," *arXiv:1908.00702*, 2019.

[102] S. Grammatico, "Dynamic Control of Agents Playing Aggregative Games with Coupling Constraints," *IEEE Transactions on Automatic Control*, vol. 62, pp. 4537–4548, Sep. 2017.

[103] J. R. Marden, H. P. Young, G. Arslan, and J. S. Shamma, "Payoff-based dynamics for multiplayer weakly acyclic games," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 373–396, 2009.

[104] T. Tatarenko and M. Kamgarpour, "Learning generalized nash equilibria in a class of convex games," *IEEE Transactions on Automatic Control*, vol. 64, pp. 1426–1439, Apr. 2019.

[105] J. R. Marden, S. D. Ruben, and L. Y. Pao, "A Model-Free Approach to Wind Farm Control Using Game Theoretic Methods," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 4, pp. 1207 – 1214, 2013.

[106] M. Zhu and S. Martínez, "Distributed Coverage Games for Energy-Aware Mobile Sensor Networks," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 1–27, 2013.

[107] R. Srikant, *The Mathematics of Internet Congestion Control*. Birkhaeuser, 2004.

[108] S. H. Low and D. E. Lapsley, "Optimization flow control. I. Basic algorithm and convergence," *IEEE/ACM Trans. on Net.*, vol. 7, pp. 861–874, Dec. 1999.

[109] F. Paganini, Z. Wang, J. C. Doyle, and S. H. Low, "Congestion control for high performance, stability, and fairness in general networks," *IEEE/ACM Transactions on Networking*, vol. 13, no. 1, pp. 43–56, 2005.

[110] G. Chen and M. Teboulle, "Convergence Analysis of a Proximal-Like Minimization Algorithm Using Bregman Functions," *SIAM Journal on Optimization*, vol. 3, no. 3, pp. 538–543, 1993.

[111] M. Mahdavi, R. Jin, and T. Yang, "Trading Regret for Efficiency: Online Convex Optimization with Long Term Constraints," *J. Mach. Learn. Res.*, vol. 13, pp. 2503 – 2528, Jan. 2012.

[112] M. Benaïm, "Dynamics of stochastic approximation algorithms," in *Séminaire de Probabilités XXXIII* (J. Azéma, M. Émery, M. Ledoux, and M. Yor, eds.), (Berlin, Heidelberg), pp. 1–68, Springer Berlin Heidelberg, 1999.

[113] B. Bharath and V. S. Borkar, "Stochastic approximation algorithms: Overview and recent trends," *Sadhana*, vol. 24, pp. 425 – 452, 1999.

[114] V. S. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint*, vol. 48. Hindustan Book Agency, 2008.

[115] M. Duflo, *Random Iterative Models.* 1st, Berlin, Heidelberg: Springer-Verlag, 1997.

[116] S. Boucheron, G. Lugosi, and P.Massart, *Concentration inequalities: A nonasymptotic theory of independence.* Oxford University Press, 2013.

[117] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science.* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2018.

[118] N. G. Mankiw, *Principle of Economics.* South-Western Cengage Learning, 5 ed., 2011.

[119] D. Evans, "The Internet of Things: How the Next Evolution of the Internet is Changing Everything," tech. rep., CISCO, April 2011.

[120] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, "Worldwide energy needs for ICT: The rise of power-aware networking," in *2nd IEEE ANTS*, pp. 1–3, Dec. 2008.

[121] R. W. Rosenthal, "A class of games possessing pure-strategy nash equilibria," *Int. J. of Game Th.*, vol. 2, pp. 65–67, Dec 1973.

[122] D. Schmeidler, "Equilibrium points of nonatomic games," *J. of Stat. Phy.*, vol. 7, pp. 295–300, Apr 1973.

[123] J. G. Wardrop, "Some Theoretical Aspects of Road Traffic Research.," *Proc. of the Inst. of Civ. Eng.*, vol. 1, no. 3, pp. 325–362, 1952.

[124] S. Fischer and B. Vöcking, "On the evolution of selfish routing," in *ESA*, (Berlin, Heidelberg), pp. 323–334, Springer Berlin Heidelberg, 2004.

[125] A. Blum, E. Even-Dar, and K. Ligett, "Routing Without Regret: On Convergence to Nash Equilibria of Regret-Minimizing Algorithms in Routing Games," in *Proc. of the 25th Ann. ACM Symp. on Princ. of Dist. Comp.*, pp. 45 – 52, ACM, 2006.

[126] W. Krichene, B. Drighés, and A. Bayen, "Online Learning of Nash Equilibria in Congestion Games," *SIAM Journal on Control and Optimization*, vol. 53, no. 2, pp. 1056–1081, 2015.

[127] T. Alpcan and T. Basar, "A game-theoretic framework for congestion control in general topology networks," in *Proc. of the 41th IEEE International Conference on Decision and Control*, vol. 2, pp. 1218–1224, 2002.

[128] A. Ozdaglar and R. Srikant, "Incentives and pricing in communication networks," in *In Algorithmic Game Theory*, pp. 571–591, Cambridge Press, 2007.

[129] F. Farokhi and K. H. Johansson, "A piecewise-constant congestion taxing policy for repeated routing games," *Transportation Research Part B: Methodological*, vol. 78, pp. 123 – 143, 2015.

[130] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Beh.*, vol. 14, pp. 124 – 143, May 1996.

[131] J. Yuan and A. Lamperski, "Online convex optimization for cumulative constraints," in *Proc. of the 32nd Int. Conf on neu. Inf. Process. Sys.*, pp. 6140 – 6149, 2018.

[132] H. Yu and M. J. Neely, "A Low Complexity Algorithm with $\mathcal{O}(\sqrt{T})$ Regret and Finite Constraint Violations for Online Convex Optimization with Long Term Constraints," *arXiv:1604.02218*, 2016.

[133] R. Jenatton, J. C. Huang, and C. Archambeau, "Adaptive Algorithms for Online Convex Optimization with Long-term Constraints," in *Proc. of the 33rd Int. Conf on Mach. Learn.*, vol. 48, pp. 402 – 411, 2016.

[134] H. Yu, M. J. Neely, and X. Wei, "Online Convex Optimization with Stochastic Constraints," in *Proc. of the 31st Int. Conf. on Neur. Inf. Process. Sys.*, pp. 1427 – 1437, 2017.

[135] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power generation, operation, and control.* Wiley-Interscience, 2014.

[136] T. Chen, A. Mokhtari, X. Wang, A. Ribeiro, and G. B. Giannakis, "Stochastic Averaging for Constrained Optimization With Application to Online Resource Allocation," *IEEE Transactions on Signal Processing*, vol. 65, pp. 3078–3093, June 2017.

[137] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1936–1948, Dec. 1992.

[138] M. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems.* Morgan & Claypool, 2010.

[139] Y. Li, G. Qu, and N. Li, "Using predictions in online optimization with switching costs: A fast algorithm and a fundamental limit," in *2018 Annual American Control Conference (ACC)*, pp. 3008–3013, June 2018.

[140] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, pp. 2941 – 2962, Nov. 2017.

[141] T. Keviczky, F. Borrelli, and G. J. Balas, "Decentralized receding horizon control for large scale dynamically decoupled systems," *Automatica*, vol. 42, no. 12, pp. 2105 – 2115, 2006.

[142] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, vol. 3 of *Found. Trends Mach. Learn.* Now Publishers Inc., Jan. 2011.

[143] L. Shu, M. Mukherjee, X. Xu, K. Wang, and X. Wu, "A survey on gas leakage source detection and boundary tracking with wireless sensor networks," *IEEE Access*, vol. 4, pp. 1700 – 1715, 2016.

[144] K. Zhou and S. I. Roumeliotis, "Multirobot active target tracking with combinations of relative observations," *IEEE Transactions on Robotics*, vol. 27, pp. 678 – 695, Aug. 2011.

[145] J. A. Bazerque and G. B. Giannakis, "Distributed spectrum sensing for cognitive radio networks by exploiting sparsity," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1847 – 1862, March 2010.

[146] P. D. Lorenzo and S. Barbarossa, "A bio-inspired swarming algorithm for decentralized access in cognitive radio," *IEEE Transactions on Signal Processing*, vol. 59, pp. 6160 – 6174, Dec. 2011.

[147] V. Gazi and K. M. Passino, *Swarm stability and optimization.* Springer-Verlag Berlin Heidelberg, 2011.

[148] J. K. Hale, *Asymptotic Behaviour of Dissipative Systems.* American Mathematical Society, 1988.

[149] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems.* Springer-Verlag Berlin Heidelberg, 2012.

[150] S. S. Sastry, *Nonlinear Systems: Analysis, Stability, and Control.* Springer-Verlag New York, 1999.

[151] U. Helmke and J. Moore, *Optimization and dynamical systems.* Communications and control engineering series, Springer-Verlag, 1994.

[152] A. Bloch, ed., *Hamiltonian and gradient flows, algorithms and control.* American Mathematical Society, 1994.

[153] A. S. Nemirovsky and D. Yudin, *Problem Complexity and Method Efficiency in Optimization.* Wiley-Interscience series in discrete mathematics, Wiley, 1983.

[154] W. Su, S. Boyd, and E. Candés, "A differential equation for modeling nesterov's accelerated gradient method: Theory and insights," in *NIPS 2014*, 2014.

[155] W. Krichene, A. Bayen, and P. Bartlett, "Accelerated mirror descent in continuous and discrete time," in *NIPS 2015*, 2015.

[156] A. Wibisono, A. C. Wilson, and M. I. Jordan, "A variational perspective on accelerated methods in optimization," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7351–E7358, 2016.

[157] W. Krichene and P. Bartlett, "Acceleration and averaging in stochastics descent dynamics," in *NIPS 2017*, 2017.

[158] H. Attouch, Z. Chbani, J. Peypouquet, and P. Redont, "Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity," *Math. Program. Ser. B*, no. 168, pp. 123 – 175, 2018.

[159] H. Attouch and A. Cabot, "Asymptotic stabilization of inertial gradient dynamics with time-dependent viscosity," *Journal of Differential Equations*, vol. 263, no. 9, pp. 5412 – 5458, 2017.

[160] J. N. Tsitsiklis, "Pproblem in decentralized decision making and computations," tech. rep., DTIC, 1984.

[161] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, 1986.

[162] B. Touri and A. Nedic, "Distributed consensus over network with noisy links," in *2009 12th International Conference on Information Fusion*, pp. 146 – 154, 2009.

[163] R. L. G. Cavalcante and S. Stanczak, "A distributed subgradient method for dynamic convex optimization problems under noisy information exchange," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 243 – 256, 2013.

[164] H. Attouch and M.-O. Czarnecki, "Asymptotic behavior of coupled dynamical systems with multiscale aspects," *Journal of Differential Equations*, vol. 248, no. 6, pp. 1315 – 1344, 2010.

[165] S. S. Kia, J. Cortés, and S. Martínez, "Distributed convex op-timization via continuous-time coordination algorithms withdiscrete-time communication," *Automatica*, vol. 55, no. C, pp. 254–264, 2015.

[166] D. Wang, D. Wang, W. Wang, Y. Liu, and F. E. Alsaa, "A modified distributed optimization method for both continuous-time and discrete-time multi-agent system," *Neurocomputing*, vol. 275, pp. 725–732, 2018.

[167] M. Raginsky and J. Bouvrie, “Continuous-time stochastic mirror descent on a network: Variance reduction, consensus, convergence,” in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 6793 – 6800, 2012.

[168] S. Pu and A. Garcia, “A flocking-based approach for distributed stochastic optimization,” *Operations Research*, vol. 66, no. 1, 2018.

[169] F. R. K. Chung, *Spectral Graph Theory*, vol. 92. American Mathematical Society, 1997.

[170] R. A. Horn and C. R. Johnson, *Matrix Analysis*. 2 ed., 2012.

[171] X. Mao, *Stochastic Differential Equations and their Applications*. Horwood Publishing Chichester, 1997.

[172] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*. Springer-Verlag New York, 1998.

[173] A. H. Sayed, S. Tu, J. Chen, X. Zhao, and Z. J. Towfic, “Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 155–171, 2013.

[174] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, “Graph signal processing: Overview, challenges, and applications,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[175] A. Sandryhaila and J. M. F. Moura, “Big Data Analysis with Signal Processing on Graphs: Representation and processing of massive data sets with irregular structure,” *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[176] M. H. DeGroot, “Reaching a consensus,” *J. Amer. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.

[177] A. De, S. Bhattacharya, P. Bhattacharya, N. Ganguly, and S. Chakrabarti, “Learning a Linear Influence Model from Transient Opinion Dynamics,” in *Proc. of the 23rd ACM Int. Conf. on Inf. and Know. Man.*, 2014.

[178] A. Das, S. Gollapudi, and K. Munagala, “Modeling Opinion Dynamics in Social Networks,” in *Proc. of the 7th ACM Int. Conf. on Web Search and Data Mining*, pp. 403–412, 2014.

[179] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo, “Opinion dynamics and the evolution of social power in influence networks,” *SIAM Review*, vol. 57, no. 3, pp. 367–397, 2015.

[180] A. G. Chandrasekhar, H. Larreguy, and J. P. Xandri, "Testing Models of Social Learning on Networks: Evidence from a Lab Experiment in the Field," Working Paper 21468, NBE, 2015.

[181] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dyn. Games Appl.*, vol. 1, no. 1, pp. 2–49, 2011.

[182] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in Multiagent Coordination, Consensus, and Flocking," in *Proc. of the 44th IEEE CDC*, pp. 2996–3000, 2005.

[183] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. of Par. and Dist. Comp.*, vol. 67, no. 1, pp. 33 – 46, 2007.

[184] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[185] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Tr. IT*, vol. 52, no. 6, pp. 2508–2530, 2006.

[186] D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi, "Spread of (mis) information in social networks," *Games and Economic Behavior*, vol. 70, no. 2, pp. 194–227, 2010.

[187] C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii, "Ergodic randomized algorithms and dynamics over networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 78–87, 2015.

[188] L. Li, A. Scaglione, A. Swami, and Q. Zhao, "Consensus, Polarization and Clustering of Opinions in Social Networks," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 31, pp. 1072–1083, 2013.

[189] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, vol. 100, no. 5, pp. 992–1026, 1992.

[190] A. Banerjee and D. Fudenberg, "Word-of-mouth learning," *Games and Economic Behavior*, vol. 46, no. 1, pp. 1–22, 2004.

[191] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar., "Bayesian learning in social networks," *TheReview of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.

[192] J. N. Tsitsiklis, "Problems in decentralized decision making and compution," tech. rep., MIT CAMBRIDGE LAB FOR INFORMATION AND DECISION SYSTEMS,, 1984.

[193] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gra-dient optimization algorithms," *IEEE transactions on automatic control*, vol. 31, no. 9, pp. 803 – 812, 1986.

[194] A. Jadbabaie, Jie Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[195] A. Nedic and A. Ozdaglar, "Distributed Subgradient Methods for Multi-Agent Optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[196] A. Olshevsky and J. N. Tsitsiklis, "onvergence speed in distributed consensus and averaging," *SIAM Journal onControl and Optimization*, vol. 48, no. 1, pp. 33–55, 2009.

[197] N. E. Friedkin and E. C. Johnsen, "Social influence networks and opinion change," *Advances in Group Processes*, vol. 16, no. 1, pp. 1–29, 1999.

[198] J. Ghaderi and R. Srikant, "Opinion Dynamics in Social Networks with Stubborn Agents," *Automatica*, vol. 50, no. 12, pp. 3209–3215, 2014.

[199] D. Bindel, J. Kleinberg, and S. Oren, "How bad is forming your own opinion?," in *IEEE 52nd Ann. Symp. on Found. of Comp. Sc.*, pp. 57–66, 2011.

[200] M. Ji, G. Ferrari-Trecate, M. Egerstedt, and A. Buffa, "Containment control in mobile networks," *IEEE Transactions on Automatic Control*, vol. 53, no. 8, 2008.

[201] C. D. Meyer, "Stochastic complementation, uncoupling markov chains, and the theory of nearly reducible systems," *SIAM Review*, vol. 31, no. 2, 1989.

[202] H. A. Simon and A. Ando, "Aggregation of Variables in dynamic systems," *Econometrica*, vol. 29, no. 2, p. 1, 1961.

[203] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences.* Society for Industrial and Applied Mathematics, 1994.

[204] Y. Nesterov, *Lectures on Convex Optimization.* Springer International Publishing, 2018.

[205] F. Salehisadaghiani and L. Pavel, "Distributed nash equilibrium seeking: A gossip-based algorithm," *Automatica*, vol. 72, pp. 209 – 216, 2016.

[206] L. A. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, pp. 39–43, 1953.

[207] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *The Journal of Mathematical Sociology*, vol. 2, no. 1, pp. 113–120, 1972.