Social Optimality via Dynamic Tolling and Adaptive Incentive Design

Kshitij Kulkarni
S. Shankar Sastry

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2022-217
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2022/EECS-2022-217.html

August 18, 2022

Acknowledgement

I thank my advisor, Professor Shankar Sastry, and Professor Alexandre Bayen for their guidance and feedback on this thesis. I would also like to thank Chinmay Maheshwari and Professor Manxi Wu, with whom the papers this thesis is based on were jointly authored.

Social Optimality via Dynamic Tolling and Adaptive Incentive Design

Kshitij Kulkarni

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

Committee:

Shankar Sastry

Professor S. Shankar Sastry
Research Advisor

8/18/22

(Date)

Professor Alexandre Bayen
Second Reader

(Date)

Social Optimality via Dynamic Tolling and Adaptive Incentive Design

by

Kshitij Kulkarni

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Master of Science

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor S. Shankar Sastry, Chair
Professor Alexandre Bayen

Summer 2022

Abstract

Social Optimality via Dynamic Tolling and Adaptive Incentive Design

by

Kshitij Kulkarni

Master of Science in Electrical Engineering and Computer Science

University of California, Berkeley

Tolls and incentives are tools that can be used by societal-scale system designers to steer selfish players to social optimality in a variety of settings. Some representative examples include traffic routing and competition between firms. These tools have been studied heavily in static settings, but in many situations, players are learning or updating their strategies in response to changing system conditions. We ask two questions: 1) In setting tolls on traffic networks, how can a traffic authority design tolls that induce socially optimal traffic loads with dynamically arriving travelers who make selfish routing decisions? 2) In the more general setting of atomic and nonatomic games, how can a planner adaptively incentivize selfish agents who are learning in a strategic environment to induce a socially optimal outcome in the long run?

To answer both questions, we propose toll and incentive updates that account for the *externality* created by the players as measured by the planner's objective function over time. These dynamics, when coupled to the strategy update dynamics of the selfish players, run at a slower timescale. In the case of traffic routing, we consider load updates in which inflows and outflows into the network are stochastically realized, and such that the travelers are myopic. We show that the toll and load updates converge to a neighborhood of the socially optimal loads. In the general case of atomic and nonatomic games, we provide sufficient conditions for the incentive and strategy updates to converge asymptotically to social optimality, and provide applications that satisfy these conditions, including Cournot competition and quadratic aggregative games. This thesis is the compilation of the two works [28] and [27].

Contents

Chapter 1

Introduction

In recent years, traffic networks have experienced increasing problems due to congestion, which occurs as a result of selfish users on the network choosing routes that are optimal for themselves without considering their impact on the network [22], [42]. A generalization of this problem manifests in many societal-scale systems in which selfish players make decisions that affect themselves and other players. The outcome arising from such strategic interactions – Nash equilibrium – often leads to suboptimal societal outcomes. This is due to the fact that individual players often ignore the externality of their actions (i.e. how their actions affect the cost of others) when minimizing their own cost. In traffic networks, tolling, when properly set in response to real-time conditions, can effectively alleviate congestion and even induce socially optimal loads [43, 33]. More generally, an important way to address the issue of externality is to provide players with incentives that align their individual goal of cost minimization with the goal of minimizing the total cost of the society ([33, 25, 20, 2]).

As an example, a traffic authority may want to minimize the total congestion on the network, or maximize the throughput of the network subject to the selfish route choices of travelers. This is also the case in more general games (for example, in firm competition), in which there is a tension between selfish users' decisions and the welfare of the collective. The static properties of such games, along with optimal tolls to alleviate congestion and other negative externalities, have been widely studied in the literature [38].

However, what distinguishes traffic networks in reality from the above discussion is that demand arrives at the network over time, and the decision to place flow on the network is influenced by the current congestion levels on the edges at every time. That is, the users on the network are dynamically making selfish decisions and placing flow on the network in response to changing network conditions. To design tolls that induces efficient traffic loads, one must account for the dynamics of incoming and outgoing traffic demands and the selfish nature of travelers' route choices. In particular, the tolls must be updated in response to changing traffic conditions.

This is also the case in general games, where oftentimes, players are *learning* to a selfish equilibrium (such as a Nash equilibrium) by deploying a class of learning updates ([1, 12, 28]). This intrinsically causes the emergence of dynamics in their strategies. Both these examples suggest the need for tolling and incentive schemes that adjust to the changing state of the traffic network, and to the changing strategies of the players over time, respectively. We therefore ask the following question:

How can *adaptive* tolling and incentive schemes be designed for situations in which *strategic* players are dynamically responding to changing conditions?

The desiderata for such schemes (specified for traffic networks as a concrete example) are proposed here as follows:

1. Stochastically arriving travelers on a traffic network should adjust their strategies (fraction of the arriving travelers to send on a particular route) according to the current flow on the network. The travelers should make the decision to send flow on the edges selfishly, by choosing routes that approximately minimize their travel time.

2. Correspondingly, the planner should use only the current network state to set tolls on the edges of the network. The tolls enter the costs incurred by the travelers on every edge.

3. The tolls should evolve at a slower timescale than the strategies. This is to ensure that travelers on traffic networks do not experience rapidly changing tolls over days.

4. The flows and tolls should jointly evolve towards a socially optimal equilibrium. That is, asymptotically, the flows converge to a game-theoretically relevant equilibrium that

is also socially optimal as measured by a social cost, and the tolls converge to the corresponding quantities that make the selfish equilibrium socially optimal.

These can be extended to incentive schemes for general atomic and nonatomic games by replacing the specific traffic dynamics with a class of learning updates for the players' strategies, and a corresponding incentive update. In general, therefore, we propose coupling slowly varying incentive updates with fast strategy updates for selfish players in games such that jointly, the strategy and incentive dynamics converge to social optimality.

The above dynamic situation is less understood than the various results on static tolling known in the literature. Motivated by this gap, this thesis is organized as follows:

- Chapter 2 analyzes a two-timescale tolling scheme coupled with a game-theoretically relevant traffic load update on a parallel link network, following [27]. The intuition behind this scheme is that for every value of the toll, stochastically arriving traffic uses a selfish route choice to place loads on the network. The loads quickly converge to a perturbed equilibrium which is a function of the tolls. As the tolls slowly vary over time and adjust given the externality on the network at every time-step, the joint load and toll update concentrate in a neighborhood of social optimality. We provide numerical experiments that validate our theoretical results. This scheme can be used by toll designers to better control congestion.

- Chapter 3 extends this framework to a wider class of game learning dynamics coupled with incentive update schemes in atomic and nonatomic games, following [28]. Here, the goal is to steer players who are learning (updating their strategies) according to a class of well-known dynamics (for example, best-response or equilibrium update) in general games to social optimality as measured by a social cost. The paper constructs an incentive update that allows players to be steered to social optimality. We present applications of the theory, including incentive design in nonatomic routing games, atomic Cournot games, and atomic networked aggregative games.

Both the tolls in the setting of traffic routing and incentives in general games have the feature that they are computed by measuring the *externality* created by the actions of selfish players relative to the social cost. Combined, these serve to provide traffic authorities, regulators, and societal-scale system designers with methods for designing incentives that adapt to dynamic changes in the state of the system and in the strategies of selfish players. The proposed incentives have the role of steering the selfish players to social optimality over time, such that the joint strategy and incentive updates converge to an equilibrium desired by the planner. The incentives have a natural economic meaning given by the externality generated by a player on others in atomic games, and by a strategy on other strategies in nonatomic games. Further, the incentives can be implemented only using locally available information (the strategy of each player, and evaluations of the derivatives of the players' cost functions and the social cost function). As societal-scale systems like traffic networks experience increased usage and issues due to congestion, the proposed schemes serve to enable tractable implementations of tolls and incentives.

Chapter 2

Dynamic Tolling for Inducing Socially Optimal Traffic Loads

In this chapter, we propose a discrete-time stochastic dynamics to capture the joint evolution of loads and tolls in a parallel traffic network. In each time step of the dynamics, non-atomic travelers arrive at the origin of the network, and they make routing decisions according to a perturbed best response based on the travel time cost and toll of each link in that step. Additionally, a fraction of load on each link leaves the network. Both the incoming and outgoing demand are randomly realized, and are identically and independently distributed across steps. Therefore, the discrete-time stochastic dynamics of loads forms a Markov process, which is governed by the total arriving demand, the stochastic user equilibrium, and the load discharge rate. Furthermore, at each time step, a traffic authority adjusts the toll on each link by interpolating between the current toll and a new increment dependent on the marginal cost of travel time given the load at that step.

In our setting, the dynamics of the toll evolves at a slower time-scale compared to that of the load dynamics. In practice, fast changing tolls are undesirable ([15]). This property ensures that the tolls change very slowly, and thus travelers can view the tolls as static when they make routing decisions at the arrival.

We show that the loads and tolls in the discrete-time stochastic dynamics asymptotically concentrate in a neighborhood of a unique fixed point with high probability. The fixed point load is *socially optimal* in that it minimizes the total travel time costs when the incoming and outgoing traffic demands reach a *steady state*, and the fixed point toll on each link equals to the marginal cost. That is, with high probability, our dynamic tolling eventually induces the socially optimal loads that accounts for the incoming and outgoing travelers and their selfish routing behavior. Furthermore, we emphasize that our dynamic tolling is *distributed* in that the traffic authority only uses the information of the cost and load on each link to update its toll.

Our technical approach to proving the main result involves: *(i)* Constructing a continuous-time deterministic dynamical system associated with the two timescale discrete-time stochastic dynamics; *(ii)* Demonstrating that the flow of the continuous-time dynamical system has a unique fixed point that corresponds to the socially optimal load and tolls; *(iii)* Proving that the unique fixed point of the flow of the continuous time dynamical system is globally stable. In particular, we apply the theory of two time-scale stochastic approximation to show that the loads and tolls under the stochastic dynamics concentrates with high probability in the neighborhood of the fixed point of the flow of continuous time dynamics constructed in *(i)* ([5]). Additionally, our proof in *(ii)* on the uniqueness and optimality of fixed point of the flow of continuous time dynamics builds on a variational inequality, and extends the analysis of stochastic user equilibrium in static routing games to account for the steady state of the network given the incoming and outgoing demand ([10]). Furthermore, we show that the continuous time dynamical system is cooperative, and thus its flow must converge to its fixed point ([19]).

Our model and results contribute to the rich literature on designing tolling mechanisms for inducing socially optimal route loads. Classical literature on static routing games has focused on measuring the inefficiency of selfish routing by bounding the "price of anarchy", and designing marginal cost tolling to induce socially optimal route loads ([9, 39, 38]). In static routing games, optimal tolling does not account for the dynamic arrival and departure of travelers. The computation of optimal tolls relies on knowledge of the entire network structure and the equilibrium route flows, which are challenging to compute.

Dynamic toll pricing has been studied in a variety of settings to account for the continuous incoming and outgoing traffic demand. The paper [3] analyzed a discrete-time stochastic dynamics of of non-atomic travelers, and discussed the impact of tolling on routing strategies. We consider an adaptive adjustment of tolls using marginal toll pricing. This allows us to analyze the long-run outcomes of the joint evolution of the route loads and tolls, and shows that the tolls eventually induce a socially optimal traffic load associated with the steady state of the network. Additionally, we prove that the monotonicity condition – the equilibrium routing strategy is monotonic in tolls (which is an assumption in [3]) – holds for any equilibrium routing strategy.

Moreover, [12] proposed a continuous-time dynamical system to study socially optimal tolling when strategic travelers continuously arrive and make selfish routing decisions. In their model, the incoming and outgoing traffic demands are equal so that the aggregate load in the network is a constant. In our setting, both the incoming and outgoing traffic demands are random variables. Therefore, our fixed point analysis needs to account for the total load at the steady state of the network. Moreover, [12] assumes that the tolls are adjusted at a faster time scale than their traveler's route preferences, while we assume that the update of tolls is at a slower timescale compared to the change of routing decisions.

Finally, this chapter is also related to the literature on learning in routing games, and learning for tolling with unknown network conditions. In particular, a variety of algorithms have been proposed to study how travelers learn an equilibrium by repeatedly adjusting their routing decisions based on the observed travel time in the network (e.g. [11, 24, 23]). Additionally, papers [15, 36, 30, 47] have analyzed how the traffic authority adaptively updates the tolls while learning the unknown network condition using crowd-sourced data on traffic load and time costs.

This chapter is organized as follows: we introduce the dynamic tolling model and the discrete time stochastic dynamics in Section 2.1. We present the main result in Section 2.2, and numerical examples in Section 2.3. We present the key ideas of our proof techniques in the main text, and include all proofs in Appendix B.

Notations

We denote the set of non-negative real numbers by $\mathbb{R}_{\geq 0}$. For any natural number R, we succinctly write $[R] = \{1, 2, \ldots, R\}$. For any vector $x \in \mathbb{R}^n$, we define $\mathsf{diag}(x) \in \mathbb{R}^{n \times n}$ to be a diagonal matrix with its diagonals filled with entries of x.

2.1 Model

Consider a parallel link network with R links connecting a single source-destination pair. At each time step $n = 1, 2, \ldots,$ a non-atomic traveler population arrives at the source node, and is routed through links R to the destination node.

At time step n, the traffic load (i.e. amount of travellers) on link $i \in [R]$ is $X_i(n)$. The latency function on any link $i \in [R]$ is a function of load on the link: $\ell_i : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$. We assume that the latency function is *strictly-increasing* and *convex*, which reflect the congestible nature of links and the fact that the latency increases faster when the load is higher.

A traffic authority sets toll prices on links, denoted as $P(n) = (P_i(n))_{i \in \mathbb{R}}$, where $P_i(n)$ is the toll on link i at step n. The *cost function* on link $i \in [R]$ at step n, denoted $c_i(X_i(n), P_i(n))$, is the sum of the latency function $\ell_i(X_i(n))$ and the charged toll price (see Fig. 2.1):

$$c_i(X_i(n), P_i(n))) \coloneqq \ell_i(X_i(n)) + P_i(n) \tag{2.1}$$

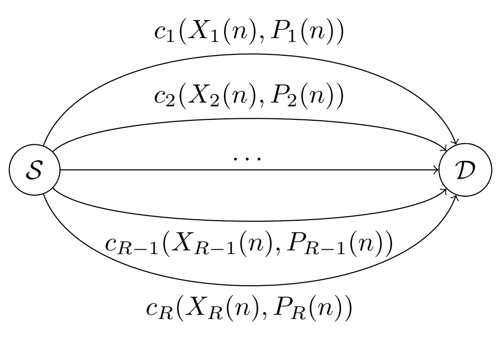


Figure 2.1: An $R-$link parallel network with source node $\mathcal{S}$, destination node $\mathcal{D}$, and cost functions at step n.

Furthermore, we define $\tilde{c}_i(X_i(n), P_i(n), \nu_i(n)) \coloneqq c_i(X_i(n), P_i(n)) + \nu_i(n)$ to be the randomly realized travel cost experienced by the travelers, where $\nu_i(n)$ is an identically and independently distributed (i.i.d) random variable with zero mean.

The demand of traffic arriving at the source node at step $n+1$ is a random variable $\zeta(n+1)$. We assume that $\{\zeta(n)\}_{n\in\mathbb{N}}$ are i.i.d with bounded support $[\underline{\lambda}, \bar{\lambda}]$ and the mean λ (i.e. $\mathbb{E}[\zeta(n)] = \lambda$ for all n). At step $n+1$, travelers make routing decisions $F^{(\mathsf{in})}(n+1) = \left(F_i^{(\mathsf{in})}(n+1)\right)_{i\in[R]}$ based on the latest cost vector at step n, where $F_i^{(\mathsf{in})}(n+1)$ is the demand of travelers who choose link i at step $n+1$. We assume that $F^{(\mathsf{in})}(n+1)$ is a perturbed best response defined as follows:

Definition 2.1.1. ***(Perturbed Best Response Strategy)*** *At any step $n+1$, the routing strategy $F_i^{(in)}(n+1)$ is* perturbed best response *if for all $i \in [R]$,*

$$F_i^{(in)}(n+1) := \frac{\exp(-\beta c_i(X_i(n), P_i(n)))}{\sum_{j\in[R]} \exp(-\beta c_j(X_j(n), P_j(n)))}\zeta(n), \tag{2.2}$$

where $\beta \in [0, \infty)$.

The perturbed best response strategy reflects the myopic nature of travelers' route choices. In particular, β is a dispersion parameter that governs the relative weight of link costs in making routing decisions. If $\beta \uparrow \infty$, $F_i^{(\mathsf{in})}(n+1)$ is a best response strategy in that travelers only take links with the minimum cost in $F_i^{(\mathsf{in})}(n+1)$. On the other hand, if $\beta \downarrow 0$, then $F_i^{(\mathsf{in})}(n+1)$ assigns the arrival demand uniformly across all links.

Furthermore, the proportion of load discharged from link $i \in [R]$ at step $n+1$ is given by the random variable $\xi_i(n+1) \in (0,1)$. We assume that $\{\xi_i(n)\}_{i\in[R],n\in\mathbb{N}}$ are i.i.d. with bounded support $[\underline{\mu}, \bar{\mu}]$ and mean μ (i.e. $\mathbb{E}[\xi_i(n)] = \mu$ for all $i \in [R]$). Thus the load discharged from link $i \in [R]$ at step $n+1$ is given by

$$F_i^{(\mathsf{out})}(n+1) := X_i(n)\xi_i(n+1). \tag{2.3}$$

In each step n, the load on each link is updated as follows:

$$X_i(n+1) = X_i(n) + F_i^{(\mathsf{in})}(n+1) - F_i^{(\mathsf{out})}(n+1). \tag{2.4}$$

We note that the stochasticity of the load update arises from the randomness in the incoming load $F^{(\mathsf{in})}(n+1)$ and the outgoing load $F^{(\mathsf{out})}(n+1)$. We define

$$h_i(x_i,p_i) \coloneqq \frac{\lambda}{\mu}\frac{\exp(-\beta c_i(x_i,p_i))}{\sum_{j\in[R]}\exp(-\beta c_j(x_j,p_j))}. \tag{2.5}$$

where $x,p\in\mathbb{R}^R$, and λ (resp. μ) is the mean of incoming (resp. outgoing) load respectively. Using (2.2), (2.3), and (2.5), we can re-write (2.4) as follows:

$$X_i(n+1) = (1-\mu)X_i(n) + \mu h_i(X_i(n),P_i(n)) + \mu M_i(n+1), \tag{Update-X}$$

where

$$\begin{aligned} M_i(n+1) \coloneqq\; & h_i(X_i(n),P_i(n))(\zeta(n+1)-\lambda) \\ & - X_i(n)\left(\xi_i(n+1)-\mu\right). \end{aligned} \tag{2.6}$$

The central authority updates the toll vector $P(n)\in\mathbb{R}^R$ at each step n as follows:

$$P_i(n+1) = (1-a)P_i(n) + aX_i(n)\frac{d\ell_i(X_i(n))}{dx} \tag{Update-P}$$

where $i\in[R], n\in\mathbb{N}$ and $a\in[0,1]$ is the step size. That is, the updated toll is an interpolation between the current toll and the *marginal cost* of the link given the current load. We note that the toll is updated in a distributed manner in that $P_i(n+1)$ only depends on the load and cost on link i. The updates of $(X(n),P(n))$ are jointly governed by the stochastic updates in (Update-X) and (Update-P). We assume that tolls are updated at a slower timescale compared with the load. That is, $a \ll \mu$, where a (resp. μ) is the step size in (Update-P) (resp. (Update-X)).

2.2 Main Results

In Section 2.2.1, we present a continuous-time deterministic dynamical system that is associated with the discrete-time updates (Update-X) and (Update-P). We show that the flow of the continuous-time dynamical system has a unique fixed point that corresponds to the perturbed socially optimal load (refer Definition 2.2.2) and the socially optimal tolling. In Section 2.2.2, we apply the two timescale stochastic approximation theory to show that $(X(n),P(n))$ in the discrete-time stochastic updates concentrate on a neighborhood of the fixed point of the flow of the continuous-time dynamical system. Therefore, our dynamical tolling eventually induces a perturbed socially optimal load vector with high probability.

2.2.1 Continuous-time dynamical system

We first introduce a *deterministic* continuous-time dynamical system that corresponds to (Update-X) – (Update-P). The time evolution in continuous-time dynamical system is denoted by t and it is related with the discrete-time step n as $t = an$, where a is the stepsize in (Update-P). We define $x(t) \in \mathbb{R}^R$ as the load vector and $p(t) \in \mathbb{R}^R$ as the toll vector at time $t \in [0, \infty)$ We also define $\epsilon \coloneqq \frac{a}{\mu}$, where μ (resp. a) is the stepsize of the discrete-time load update (resp. toll update). Since the toll update occurs at a slower timescale compared to the load update (i.e. $a \ll \mu$), we have $\epsilon \ll 1$.

The continuous-time dynamical system is as follows:

$$\begin{aligned} \dot{x}_i(t) &= \frac{1}{\epsilon}\left(h_i(x(t), p(t)) - x_i(t)\right), \\ \dot{p}_i(t) &= -p_i(t) + x_i(t)\frac{d\ell_i(x_i(t))}{dx}, \quad \forall\, t \geq 0. \end{aligned} \tag{2.7}$$

We introduce the following two definitions:

Definition 2.2.1. ***(Stochastic user equilibrium)*** *For any fixed $p \in \mathbb{R}^R$, a load vector $\bar{x}^{(\beta)}(p)$*[1] *is the stochastic user equilibrium corresponding to the toll vector p and demand $\frac{\lambda}{\mu}$ if for all $i \in [R]$:*

$$\bar{x}_i^{(\beta)}(p) = \frac{\lambda}{\mu}\frac{\exp(-\beta c_i(\bar{x}_i^{(\beta)}(p), p_i))}{\sum_{j\in[R]}\exp(-\beta c_j(\bar{x}_j^{(\beta)}(p), p_j))}. \tag{2.8}$$

Given the stochastic user equilibrium in (2.8), all travelers with total demand of λ/μ make routing decisions in the network according to a perturbed best response given the latency functions on links and tolls p. We note that demand λ/μ is the expected value of the total demand in network at step n when $n \uparrow \infty$. This is because in each step n, the expected value of the total demand in network is $\sum_{m=1}^{n}(1-\mu)^{n-m}\lambda$, where $(1-\mu)^{n-m}\lambda$ is the expected incoming demand in step m that remains in the network in step n. Thus, as $n \uparrow \infty$, the expected value of the total demand is λ/μ.

Definition 2.2.2. ***(Perturbed socially optimal load)*** *A load vector $\bar{y}^{(\beta)}$ is a perturbed*

[1] We explicitly state the dependence on p, β in order relate the definition to later results.

socially optimal load if $\bar{y}^{(\beta)}$ *minimizes the following convex optimization problem:*

$$\begin{aligned} \min_{y\in\mathbb{R}^R} \quad & \sum_{i\in[R]} y_i \ell_i(y_i) + \frac{1}{\beta}\sum_{i\in[R]} y_i \log(y_i) \\ s.t \quad & \sum_{i\in[R]} y_i = \frac{\lambda}{\mu}. \end{aligned} \tag{2.9}$$

A commonly used notion to quantify the social objective is the total latency experienced on the network [38, 15]. Note that (2.9) is an *entropy* regularized social objective function where the regularization weight depends on β. As $\beta \uparrow \infty$, the perturbed socially optimal load $\bar{y}^{(\beta)}$ becomes the socially optimal flow, which minimizes the total latency.

The following theorem shows that the flow of the continuous-time dynamical system has a unique fixed point, and this fixed point corresponds to the perturbed socially optimal load.

Theorem 2.2.3. *The flow of the continuous-time dynamical system* (2.7) *has a* unique *fixed point* $(\bar{x}^{(\beta)}(\bar{p}), \bar{p})$ *such that* $\bar{x}^{(\beta)}(\bar{p})$ *is a stochastic user equilibrium corresponding to* $\bar{p}$*, and*

$$\bar{p}_i = \bar{x}_i^{(\beta)}(\bar{p}) \frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx}. \tag{2.10}$$

Moreover, $\bar{x}^{(\beta)}(\bar{p})$ *is the perturbed socially optimal load.*

At $(\bar{x}_i^{(\beta)}(\bar{p}), \bar{p})$, the load on the network is a stochastic user equilibrium given $\bar{p}$ and demand λ/μ. This implies that travelers' routing strategy, when averaged over all time steps, is a perturbed best response given $\bar{p}$ and λ/μ. Moreover, the unique price vector $\bar{p}$ is the marginal latency cost, which ensures that the stochastic user equilibrium $\bar{x}^{(\beta)}(\bar{p})$ is also a perturbed socially optimal load. As $\beta \to \infty$, $\bar{x}^{(\beta)}(\bar{p})$ becomes a socially optimal load.

We prove Theorem 2.2.3 in three parts: Firstly, we show that for any toll vector $p \in \mathbb{R}^R$, the stochastic user equilibrium $\bar{x}^{(\beta)}(p)$ exists and is unique, and can be solved as the optimal solution of a convex optimization problem (Lemma 2.2.4). Secondly, we show that there exists a unique toll vector $\bar{p}$ that satisfies (2.10) (Lemma 2.2.6). This requires us to prove that the stochastic user equilibrium $\bar{x}^{(\beta)}(p)$ is monotonic in p (Lemma 2.2.5). Finally, we conclude the theorem by proving that at $\bar{p}$ the stochastic user equilibrium $\bar{x}^{(\beta)}(\bar{p})$ is the perturbed socially optimal load (Lemma 2.2.7).

We now present the lemmas that are referred in each of the parts above, and provide the proof ideas of these results. We include the formal proofs in Appendix B.

Part 1: For any p, the load vector $\bar{x}^{(\beta)}(p)$ is the unique stochastic user equilibrium.

Lemma 2.2.4. *For every $p \in \mathbb{R}^R$, $\bar{x}^{(\beta)}(p)$ is the unique optimal solution to the following convex optimization problem:*

$$\begin{aligned} \min_{y \in \mathbb{R}^R} \quad & \sum_{i \in [R]} \int_0^{y_i} c_i(s, p_i) ds + \frac{1}{\beta} \sum_{i \in [R]} y_i \ln y_i, \\ s.t \quad & \sum_{i \in [R]} y_i = \frac{\lambda}{\mu} \end{aligned} \tag{2.11}$$

The proof of Lemma 2.2.4 follows by verifying that $\bar{x}^{(\beta)}(p)$ (refer Definition 2.2.1) satisfies the Karush–Kuhn–Tucker (KKT) conditions corresponding to (2.11), which is strictly convex problem and therefore has unique solution. We note that as $\beta \to \infty$, the stochastic user equilibrium $\bar{x}^{(\beta)}(p)$ becomes a Wardrop equilibrium, where travelers only take routes with the minimum cost.

Part 2: We first show that the stochastic user equilibrium $\bar{x}^{(\beta)}(p)$ is monotonic in p. This allows us to prove the existence and uniqueness of $\bar{p}$ given by (2.10).

Lemma 2.2.5. *(**Monotonicity of** $\bar{x}^{(\beta)}(p)$) For any $p, p' \in \mathbb{R}^R$ we have*

$$\left\langle \bar{x}^{(\beta)}(p) - \bar{x}^{(\beta)}(p'), p - p' \right\rangle < 0.$$

Furthermore, $\bar{x}^{(\beta)}(p)$ is a continuously differentiable function. Consequently, $\frac{\partial \bar{x}_i^{(\beta)}(p)}{\partial p_i} < 0$ for all $i \in [R]$.

The proof of Lemma 2.2.5 is proved by using the variational inequalities of stochastic user equilibrium, which are the first-order optimality conditions associated with (2.11). Moreover, the monotonicity property relies on the fact that the latency on each link increases in the load.

Lemma 2.2.6. *(**Existence and uniqueness of** $\bar{p}$) The price vector $\bar{p}$, defined in* (2.10), *exists and is unique.*

In Lemma 2.2.6 the existence is proved by using the Brouwer's fixed point theorem. We prove uniqueness by contradiction. More formally, let $p, p' \in \mathbb{R}^R$ be two distinct toll vectors that satisfy (2.10). Then, from (2.10), we know that for every $i \in [R]$:

$$\begin{aligned}\bar{p}_i - \bar{p}'_i = \left(\bar{x}_i^{(\beta)}(p) - \bar{x}_i^{(\beta)}(p')\right) \frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} \\ + \bar{x}_i^{(\beta)}(p') \left(\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} - \frac{d\ell_i(\bar{x}_i^{(\beta)}(p'))}{dx} \right).\end{aligned} \tag{2.12}$$

By multiplying both sides of (2.12) by $(\bar{x}_i^{(\beta)}(p) - \bar{x}_i^{(\beta)}(p'))$,and summing over all $i \in [R]$, we have

$$\left\langle \bar{x}^{(\beta)}(p) - \bar{x}^{(\beta)}(p'), p - p' \right\rangle \geq 0,$$

which contradicts the monotonicity condition proved in Lemma 2.2.5. Therefore, the price $\bar{p}$ is unique.

Part 3: Finally, we prove that given $\bar{p}$, $\bar{x}^{(\beta)}(\bar{p})$ is a perturbed socially optimal load.

Lemma 2.2.7. *The load vector $\bar{x}^{(\beta)}(\bar{p})$ is the perturbed socially optimal load.*

We prove Lemma 2.2.7 by showing that the variational inequality satisfied by the stochastic user equilibrium at $\bar{p}$ is identical to that satisfied by the perturbed socially optimal load. Lemmas 2.2.4 – 2.2.7 conclude Theorem 2.2.3.

2.2.2 Convergence

In this section, we show that $(X_i(n), P_i(n))$ induced by the discrete-time stochastic update converges to a neighborhood of $(\bar{x}^{(\beta)}(\bar{p}), \bar{p})$. The size of the neighborhood depends on the load update stepsize μ and the stepsize ratio between the two dynamics a/μ. That is, the tolls eventually induce a perturbed socially optimal flow.

Theorem 2.2.8.

$$\limsup_{n \longrightarrow +\infty} \mathbb{E}[\|X_n - \bar{x}^{(\beta)}(\bar{p})\|^2 + \|P_n - \bar{p}\|^2] = \mathcal{O}\left(\mu + \frac{a}{\mu}\right). \tag{2.13}$$

Moreover, for any $\delta > 0$:

$$\limsup_{n \longrightarrow +\infty} \Pr(\|X_n - \bar{x}^{(\beta)}(\bar{p})\|^2 + \|P_n - \bar{p}\|^2 \geq \delta) \leq \mathcal{O}\left(\frac{\mu}{\delta} + \frac{a}{\delta\mu}\right). \tag{2.14}$$

In Theorem 2.2.8, (2.13) provides the neighborhood of the socially optimal load and tolls, where the discrete-time stochastic updates converge to in expectation. In particular, as the step size μ and the stepsize ratio between the two updates $\epsilon = a/\mu$ decreases (i.e. the discrete-time step for load update is small, and the toll update is much slower than the load update), the expected value of the distance between $(X(n), P(n))$ and $(\bar{x}^{(\beta)}(\bar{p}), \bar{p})$ decreases for $n \to \infty$. Moreover, by applying Markov's inequality, (2.13) also implies that for any neighborhood of $(\bar{x}^{(\beta)}(\bar{p}), \bar{p})$, $(X(n), P(n))$ converges to that neighborhood with high probability, and this probability increases as μ and ϵ decreases.

To prove Theorem 2.2.8, we need to prove the following technical lemma:

Lemma 2.2.9.

(C1) *For all* $i \in [R]$, $\{M_i(n+1)\}_n$ *in* (2.6) *is a martingale difference sequence with respect to the filtration*

$$\mathcal{F}_n = \sigma(\cup_{i\in[R]} (X_i(1), \zeta(1), \xi_i(1), \dots, X_i(n), \zeta(n), \xi_i(n)))$$

(C2) *For all* $i \in [R]$, $X_i(n) \in \left(0, X_i(0) + \frac{\bar{\lambda}}{\underline{\mu}}\right)$. *Consequently,* $\mathbb{E}[\|X(n)\|^2] < +\infty$, $\mathbb{E}[\|P(n)\|^2] < +\infty$.

(C3) *There exists* $K > 0$ *such that for any* $i \in [R]$ *and any* n, $\mathbb{E}[|M_i(n+1)|^2 | \mathcal{F}_n] \leq K\left(1 + \|X_i(n)\|^2\right) < +\infty$.

(C4) *For any* p, *any* $\tilde{x}(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^R$ *induced by following continuous-time dynamical system*

$$\dot{\tilde{x}}_i(t) = h_i(\tilde{x}(t), p) - \tilde{x}_i(t), \qquad \forall i \in [R], \forall\ t \geq 0, \tag{2.15}$$

satisfies $\lim_{t\to\infty} \tilde{x}(t) = \bar{x}^{(\beta)}(p)$. *Futhermore,* $\bar{x}^{(\beta)}(p)$ *is Lipschitz.*

(C5) *Any* $\tilde{p}(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^R$ *induced by following continuous-time dynamical system*

$$\dot{\tilde{p}}_i(t) = -\tilde{p}_i(t) + \bar{x}_i^{(\beta)}(\tilde{p}(t)) \frac{d\ell_i(\bar{x}_i^{(\beta)}(\tilde{p}(t)))}{dx}, \quad t \geq 0, \tag{2.16}$$

satisfies $\lim_{t\to\infty} \tilde{p}(t) = \bar{p}$.

In Lemma 2.2.9, condition **(C1)** relies on the fact that both the incoming and outgoing loads are i.i.d.. Condition **(C2)** ensures the boundedness of the loads and the tolls in the discrete-time stochastic updates, and it relies on the boundedness of ζ_n and ξ_n. Condition **(C3)** ensures the boundedness of the martingale sequence $\{M_i(n+1)\}_n$, and it is built on condition **(C1)** and **(C2)**.

In conditions **(C4)**, the continuous-time dynamical system (2.15) is associated with the discrete-time load update (Update-X) when the toll is set as a constant p. That is, due to the fact that toll update is at a slower timescale compared with the load update ($\epsilon \ll 1$), the continuous-time dynamical system of load evolves as if the toll is a constant p. We prove condition **(C4)** by showing that (2.15) is cooperative (see Theorem A.0.2 in the appendix). Condition **(C4)** ensures that the load of the continuous-time dynamical system converges to the stochastic user equilibrium (Definition 2.2.1) with respect to p. Recall from Lemma 2.2.4, the stochastic user equilibrium is unique.

On the other hand, in condition **(C5)**, (2.16) is associated with (Update-P) when the load – which is updated at a faster timescale – has already converged to the stochastic user equilibrium with respect to $p(t)$. Similar to condition **(C4)**, we show that (2.16) is a cooperative dynamical system. The proof of this condition is built on the monotonicity of stochastic user equilibrium with respect to the toll (Lemma 2.2.5) and the uniqueness of $\bar{p}$ (Lemma 2.2.6). The proofs of **(C1)** – **(C5)** are in Appendix B. Based on Lemma 2.2.9, we can apply the theory of two timescale stochastic approximation with constant stepsizes:

Lemma 2.2.10 ([5])**.** *Given $\epsilon \ll 1$, (2.13) is satisfied under **(C1)** – **(C5)**.*

Lemmas 2.2.9 and 2.2.10 proves (2.13). Additionally, (2.14) can be directly derived from (2.13) using Markov's inequality. Thus, we can conclude Theorem 2.2.8.

2.3 Numerical Experiments

In this section we present numerical experiments for the results presented in Section 2.2. We observe that loads and tolls concentrate on a neighborhood of the socially optimal loads and tolls, and the continuous-time dynamical system (2.7) closely approximates the discrete-time stochastic updates (Update-X) – (Update-P). Our numerical results are consistent with Theorems 2.2.3 and 2.2.8.

Consider a network with six parallel links (i.e. $R = 6$) with the following link latency functions:

$$\ell_i(x) = ix^2 + i, \qquad \forall\, i \in [R]. \tag{2.17}$$

We set the total time steps of discrete-time stochastic update as $N = 2000$, and the dispersion parameter $\beta = 100$. We conduct four sets of experiments with the following parameters:

(**S1**) $\lambda = 0.1$, $\mu = 0.05$ and $a = 0.0015$;

(**S2**) $\lambda = 0.2$, $\mu = 0.05$ and $a = 0.0015$;

(**S3**) $\lambda = 0.1$, $\mu = 0.05$ and $a = 0$;

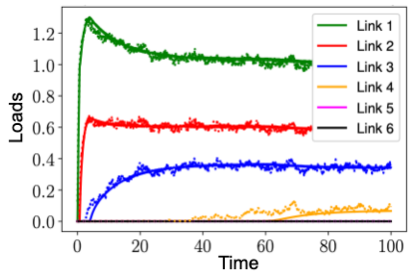
(**S4**) $\lambda = 0.2$, $\mu = 0.05$ and $a = 0$.

We note that tolls are updated with positive stepsizes in (**S1**) – (**S2**), but remain zero in (**S3**) – (**S4**). Also, the mean incoming load of travelers in each step λ is high in (**S1**) and (**S3**), and low in (**S2**) – (**S4**). In Fig. 2.2, we demonstrate the loads and tolls obtained in the discrete-time stochastic updates (Update-X) – (Update-P) (represented by dots), and those induced by the continuous-time dynamical system (2.7) (represented by solid lines) in (**S1**) and (**S2**), respectively. In Fig. 2.3, we demonstrate the loads in the discrete-time updates and the socially optimal load for (**S3**) and (**S4**). We omit the figures for tolls since tolls are not updated with step size $a = 0$.

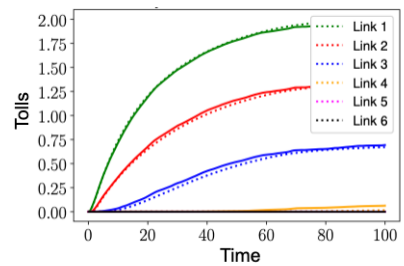
We now present the main observations from the numerical study. Firstly, we observe that in Fig. 2.2, the trajectories of continuous-time dynamical system (2.7) closely track the discrete-time stochastic load update. Moreover, we observe that it takes more time steps for the tolls (refer Fig.2.2b, Fig.2.2d) to converge compared to the loads. This is because tolls are updated at a slower timescale (i.e. $a \ll \mu$).

Secondly, we show that the repeatedly updated tolls eventually induce the perturbed socially optimal load in Fig. 2.2a and 2.2c. On the other hand, when tolls are zero (i.e. inactive) in all steps, loads converge to a stochastic user equilibrium (Fig. 2.3a and 2.3b), which is different from the socially optimal load. This means that links are inefficiently utilized when tolls are inactive. We observe that the low cost links (links 1, 2, and 3) are disproportionately over-utilized while the remaining links are underused especially in the setting with high incoming load λ (Fig.2.3b).

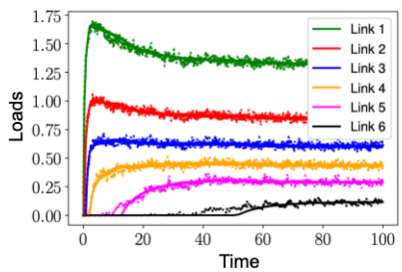
Thirdly, we note that both the perturbed socially optimal load and the stochastic user equilibrium change with the average incoming load λ. In particular, more links are utilized at fixed point with high λ (Fig. 2.2c for **(S2)** and Fig. 2.3b for **(S4)**) compared to that with low λ (Fig. 2.2a for **(S2)** and Fig. 2.3a for **(S4)**).



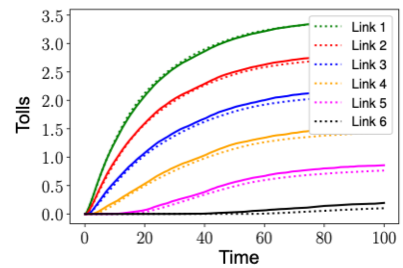
(a) Loads in (**S1**)



(b) Tolls in (**S1**)

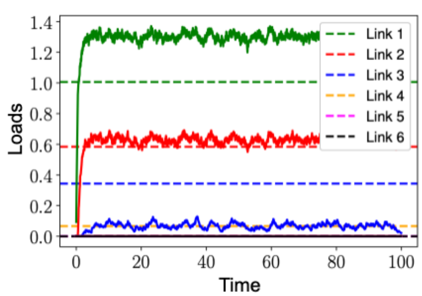


(c) Loads in (**S2**)

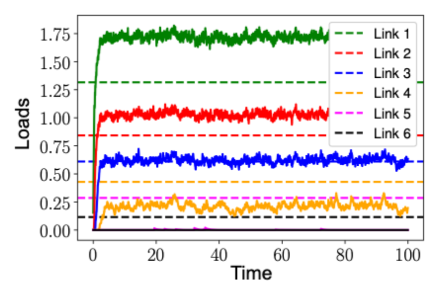


(d) Tolls in (**S2**)

Figure 2.2: Loads and Tolls in discrete-time stochastic update (Update-X) – (Update-P) (dots) and continuous-time dynamical system (2.7) (solid lines) in **(S1)** and **(S2)**.



(a) Loads in (**S3**)



(b) Loads in (**S4**)

Figure 2.3: Comparison between loads induced by the discrete-time stochastic update (Update-X)(solid lines) when $P_n = 0$ and the socially optimal load (dashed lines) in **(S3)** and **(S4)**.

Chapter 3

Inducing Social Optimality in Games via Adaptive Incentive Design

In this chapter we propose a discrete-time learning dynamics that jointly captures the players' strategy updates and the designer's updates of incentive mechanisms. Our learning dynamics can be used for both atomic games and non-atomic games. The incentive mechanism designed by the social planner sets a payment (tax or subsidy) for each player that is added to their cost function in the game. In each time step, players update their strategies based on the opponents' strategies and the incentive mechanism in the current step, and the social planner updates the incentive mechanism in response to players' current strategies. We assume that the incentive update proceeds at a slower timescale than the strategy update of players. The slower evolution of incentives is in-fact a desirable characteristic for any societal scale system, where frequent changes of incentives may lead to instability in the system and may hamper participation by players. The slow evolution of incentives allows players to consider the incentives as static while updating their strategies.

A key feature of our learning dynamics is that the incentive update in each time step is based on the externality created by each player with their current strategy. In particular, given any strategy profile, the externality of each player is evaluated as the difference between the marginal cost of their strategy on themselves and the marginal social cost. In a static incentive design problem, when all players are charged with their externality, the change of their total cost – original cost in game plus the payment – with respect to their strategy becomes identical to the change of social cost. Consequently, the induced Nash equilibrium is also socially optimal [46, 13, 35]. In our learning dynamics, the social planner accounts for the externality of each player evaluated at their current strategy, which evolves with players' strategy updates.

The externality-based incentive updates distinguish our adaptive incentive design from other recent studies on incentive mechanisms with learning agents. The paper [37] studies the problem of incentive design while learning the cost functions of players. The authors assume that both the cost functions and incentive policies are linearly parameterized, and the incentive updates rely on the knowledge of players' strategy update rules instead of just the current strategy as in our setting. Additionally, the paper [26] considers a two-timescale discrete-time learning dynamics, where players adopt a mirror descent-based strategy update, and the social planner updates an incentive parameter according to a gradient descent method. The convergence of such gradient-based learning dynamics relies on the assumption that the social cost given players' equilibrium strategy is convex in the incentive parameter. However, the convexity assumption can be restrictive since the equilibrium strategy as a function of the incentive parameter is nonconvex even in simple games.

We show that our externality-based incentive updates ensure that any fixed point of our learning dynamics corresponds to a optimal incentive mechanism, such that the induced Nash equilibrium of the game is also socially optimal (Proposition 3.2.1). This result is built on the fact that at any fixed point of our learning dynamics, the strategy profile is a Nash equilibrium corresponding to the incentive mechanism, and each player's payment equals to the externality created by their equilibrium strategy. Therefore, the equilibrium strategy associated with this externality-based payment also minimizes the social cost. Additionally, we present the sufficient conditions on the game such that the fixed point set is a singleton set, and thus the socially optimal incentive mechanism is unique (Proposition 3.2.2).

Furthermore, we provide sufficient conditions on games that guarantee the convergence of strategies and incentives induced by our learning dynamics (Theorem 3.2.3). Since the convergent strategy profile and incentive mechanism corresponds to a fixed point that is also socially optimal, these sufficient conditions guarantee that the adaptive incentive mechanism eventually induces a socially optimal outcome in the long run.

In the proof of our convergence theorem, we exploit the timescale separation between the strategy update and the incentive updates. We use tools from the theory of two-timescale dynamical systems [4] to analyze the convergence of strategy updates and incentive updates separately after accounting of time separation. In particular, the convergence of strategy updates can be derived from the rich literature of learning in games ([16],[40],[31], etc.) since the incentive mechanism can be viewed as static in the strategy updates thanks to the time separation. On the other hand, the convergence of incentive vectors can be analyzed via the associated continuous-time dynamical system, in which the value of the externality function is evaluated at the converged value of fast strategy update, which is the Nash equilibrium. Our sufficient conditions are based broadly on two main techniques of proving global stability of non-linear dynamical system: *(i)* cooperative dynamical systems theory [19] and *(ii)* Lyapunov based methods [41].

Finally, we apply our general results to three classes of games: (i) atomic networked quadratic games; (ii) atomic cournot competition; (iii) nonatomic routing games. In each class of games, we present the adaptive incentive design based on the externality of players' strategies. We also provide sufficient conditions on the game parameters and social cost functions under which the adaptive incentive design eventually induces a socially optimal outcome.

The chapter is organized as follows: in Section 3.1 we describe the setup of both atomic and non-atomic game considered here. In addition, we also provide the joint strategy and incentive update considered in this paper. We present the main results in Section 3.2 and the applications of those results in three class of games in Section 3.3.

Notations

For any vector $x \in \mathbb{R}^n$, we use x_j or x^j to denote the $j-$th component of that vector. Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we use $D_{x_i} f(x)$ to denote $\frac{\partial f}{\partial x_i}(x)$, the derivative of f with respect to x_i for any $i \in \{1, 2, ..., n\}$. For any matrix $A \in \mathbb{R}^{n\times n}$ we denote the set of eigenvalues of A

by $\mathsf{spec}(A)$. For any set A we use $\mathsf{conv}(A)$ to denote the convex hull of the set. We use k to denote the discrete-time index and t to denote the continuous-time index.

3.1 Model

We introduce both atomic and non-atomic static games in Section 3.1.1. In Section 3.1.2, we present the two-timescale dynamics of strategy learning and incentive design.

3.1.1 Static games

Atomic Games

Consider a game G with a finite set of players $\mathcal{I}$. The strategy of each player $i \in \mathcal{I}$ is $x_i \in X_i$, where X_i is a non-empty and closed interval in $\mathbb{R}$. The strategy profile of all players is $x = (x_i)_{i \in \mathcal{I}}$, and the set of all strategy profiles is $X \coloneqq \prod_{i \in \mathcal{I}} X_i$. The cost function of each player $i \in \mathcal{I}$ is $\ell_i : X \to \mathbb{R}$.[1] For any $x_{-i} = (x_j)_{j \in \mathcal{I} \setminus \{i\}}$, we assume that the cost function $\ell_i(x_i, x_{-i})$ is twice continuously differentiable and strictly convex in x_i for all $i \in \mathcal{I}$.

A *social planner* designs incentives by setting a payment $p_i x_i$ for each player i that is linear in their strategy x_i.[2] Here, p_i represents the marginal payment for every unit increase in strategy of player i. The value of p_i can either be negative or positive, which represents a marginal subsidy or a marginal tax, respectively.

Given the incentive vector $p = (p_i)_{i \in I}$, the total cost of each player $i \in \mathcal{I}$ is:

$$c_i(x, p) = \ell_i(x) + p_i x_i, \quad \forall\, x \in X. \tag{3.1}$$

A strategy profile $x^*(p) \in X$ is a *Nash equilibrium* in the atomic game G with the incentive vector p if

$$c_i(x_i^*(p), x_{-i}^*(p), p) \leq c_i(x_i, x_{-i}^*(p), p), \ \forall\, x_i \in X_i, \ \forall i \in \mathcal{I}.$$

Recall that the cost $\ell_i(x_i, x_{-i})$ is a continuous function, and is strictly convex in x_i. Additionally, the strategy set X_i is convex for each player i. Therefore, we know that Nash

[1] We measure the outcome of our games by costs instead of utilities. Equivalently, the utility of each player is the negative value of the cost.

[2] Considering a linear payment is sufficient to ensure optimal incentive design in atomic games.

equilibrium must exist and must be unique in G. Moreover, we can equivalently represent a Nash equilibrium x^* as a strategy profile that satisfies the following variational inequality ([14]):

$$\langle J(x^*(p), p), x - x^*(p) \rangle \geq 0, \quad \forall\ x \in X, \tag{3.2}$$

where $J(x^*(p), p) = (J_i(x^*(p), p))_{i \in \mathcal{I}}$, and

$$J_i(x^*(p), p) = D_{x_i} c_i(x^*(p), p) = D_{x_i} \ell_i(x^*) + p_i. \tag{3.3}$$

Furthermore, a strategy profile $x^\dagger \in X$ is *socially optimal* if $x^\dagger$ minimizes the social cost function $\Phi : X \to \mathbb{R}$. We assume that the social cost function $\Phi(x)$ is strictly convex and twice continuously differentiable in x. Then, the optimal strategy profile $x^\dagger$ is unique. Additionally, from the first order conditions of optimality, we know that $x^\dagger$ minimizes the social cost function Φ if and only if:

$$\langle \nabla \Phi(x^\dagger), x - x^\dagger \rangle \geq 0, \quad \forall\ x \in X. \tag{3.4}$$

Finally, given a strategy profile $x \in X$, we define the *externality* caused by player i as the difference between the marginal social cost, and the marginal cost of player i with respect to x_i. That is,

$$e_i(x) = D_{x_i} \Phi(x) - D_{x_i} \ell_i(x). \tag{3.5}$$

Non-atomic Games

Consider a game $\tilde{G}$ with a finite set of player populations $\tilde{\mathcal{I}}$. Each population $i \in \tilde{\mathcal{I}}$ is comprised of a continuum set of players with mass $M_i > 0$. Individual players in each population can choose an action in a finite set S_i. The strategy of population $i \in \tilde{\mathcal{I}}$ is $\tilde{x}_i = \left(\tilde{x}_i^j\right)_{j \in S_i}$, where $\tilde{x}_i^j$ is the fraction of individuals in population i who choose action $j \in S_i$. Then, the strategy set of population i is $\tilde{X}_i = \left\{\tilde{x}_i | \sum_{j \in S_i} \tilde{x}_i^j = M_i,\ \tilde{x}_i^j \geq 0, \forall j \in S_i\right\}$. The strategy profile of all populations is $\tilde{x} = (\tilde{x}_i)_{i \in \tilde{\mathcal{I}}} \in \tilde{X} = \prod_{i \in \tilde{\mathcal{I}}} \tilde{X}_i$. Given a strategy profile $\tilde{x} \in \tilde{X}$, the cost of players in population $i \in \tilde{\mathcal{I}}$ for choosing action $j \in S_i$ is $\tilde{\ell}_i^j(\tilde{x})$ which is assumed to be continuously differentiable. We denote $\tilde{\ell}_i(\tilde{x}) = (\tilde{\ell}_i^j(\tilde{x}))_{j \in S_i}$ as the vector of costs for each population $i \in \tilde{\mathcal{I}}$.

Given any $\tilde{x} \in \tilde{X}$, a social planner designs incentives by setting a payment $\tilde{p}_i^j$ for players in population i who choose action j. Consequently, given the incentive vector $\tilde{p} = \left(\tilde{p}_i^j\right)_{j \in S_i, i \in \tilde{\mathcal{I}}}$, the total cost of players in each population $i \in \tilde{\mathcal{I}}$ for choosing action $j \in S_i$ is given by:

$$\tilde{c}_i^j(\tilde{x}, \tilde{p}) = \tilde{\ell}_i^j(\tilde{x}) + \tilde{p}_i^j \quad \forall\, \tilde{x} \in \tilde{X}. \tag{3.6}$$

A strategy profile $\tilde{x}^*(\tilde{p}) \in \tilde{X}$ is a Nash equilibrium in the nonatomic game $\tilde{G}$ with $\tilde{p}$ if

$$\begin{aligned} &\forall i \in \tilde{\mathcal{I}},\ \forall j \in S_i,\ \tilde{x}_i^{j*}(\tilde{p}) > 0, \quad \Rightarrow \\ &\qquad\qquad \tilde{c}_i^j(\tilde{x}^*(\tilde{p}), \tilde{p}) \leq \tilde{c}_i^{j'}(\tilde{x}^*(\tilde{p}), \tilde{p}), \quad \forall j' \in S_i. \end{aligned}$$

Similar to atomic games, we can equivalently represent the Nash equilibrium $\tilde{x}^*(\tilde{p})$ in nonatomic game $\tilde{G}$ as a strategy profile that satisfies the following variational inequality ([40]):

$$\langle \tilde{c}(\tilde{x}^*(\tilde{p}), \tilde{p}), \tilde{x} - \tilde{x}^*(\tilde{p}) \rangle \geq 0 \quad \forall\, \tilde{x} \in \tilde{X}, \tag{3.7}$$

where $\tilde{c}(\tilde{x}^*(\tilde{p}), \tilde{p}) = (\tilde{c}_i(\tilde{x}^*(\tilde{p}), \tilde{p}))_{i \in \tilde{\mathcal{I}}}$.

Note that a Nash equilibrium always exists in a population game $\tilde{G}$ [40, Theorem 2.1.1]. Under the assumption that the cost function $\tilde{c}(\tilde{x}, \tilde{p})$ is strictly monotone in $\tilde{x}$ (Assumption 3.1.1), Nash equilibrium $\tilde{x}^*$ is also unique [40].

Assumption 3.1.1. *For every incentive vector $\tilde{p}$,*

$$\langle \tilde{c}(x, \tilde{p}) - \tilde{c}(x', \tilde{p}), x - x' \rangle > 0, \quad \forall x \neq x' \in \tilde{X}.$$

Analogous to the atomic games, a strategy profile $\tilde{x}^\dagger \in \tilde{X}$ is socially optimal if $\tilde{x}^\dagger$ minimizes a social cost function $\tilde{\Phi} : \tilde{X} \to \mathbb{R}$. We assume that $\tilde{\Phi}(\tilde{x})$ is strictly convex, and twice continuously differentiable in $\tilde{x}$. Therefore, $\tilde{x}^\dagger$ is unique, and satisfies the following variational inequality constraints:

$$\langle \nabla \tilde{\Phi}(\tilde{x}^\dagger), \tilde{x} - \tilde{x}^\dagger \rangle \geq 0, \quad \forall\, \tilde{x} \in \tilde{X}. \tag{3.8}$$

Finally, give any $\tilde{x} \in \tilde{X}$, we define the externality caused by players in population i who play action $j \in S_i$ as the difference between the marginal social cost, and the cost experienced by the players in population i who chooses action j, i.e.

$$\tilde{e}_i^j(\tilde{x}) = D_{\tilde{x}_i^j} \tilde{\Phi}(\tilde{x}) - \tilde{\ell}_i^j(\tilde{x}). \tag{3.9}$$

3.1.2 Learning dynamics

We now introduce the discrete-time learning dynamics considered in this paper. For every time step $k = 1, 2, ...$, we denote the strategy profile in the atomic game G (resp. non-atomic game $\tilde{G}$) as $x_k = (x_{i,k})_{i\in\mathcal{I}}$ (resp. $\tilde{x}_k = (\tilde{x}_{i,k})_{i\in\tilde{\mathcal{I}}}$), where $x_{i,k}$ (resp. $\tilde{x}_{i,k}$) is the strategy of player i (population i) in step k. Additionally, we denote the incentive vector as $p_k = (p_{i,k})_{i\in\mathcal{I}}$ (resp. $\tilde{p}_k = (\tilde{p}^j_{i,k})_{j\in S_i, i\in\tilde{\mathcal{I}}}$). The strategy updates and the incentive updates are presented below:

Strategy update. In each step $k+1$, the updated strategy is a linear combination of the previous strategy in stage k (i.e. x_k in G and $\tilde{x}_k$ in $\tilde{G}$), and a new strategy (i.e. $f(x_k, p_k) \in X$ in G and $\tilde{f}(\tilde{x}_k, \tilde{p}_k) \in \tilde{X}$ in $\tilde{G}$) that depends on the previous strategy and the incentive vector in stage k. The relative weight in the linear combination is determined by the step-size $\gamma_k \in (0, 1)$.

$$x_{k+1} = (1 - \gamma_k)x_k + \gamma_k f(x_k, p_k) \tag{x-update}$$

$$\tilde{x}_{k+1} = (1 - \gamma_k)\tilde{x}_k + \gamma_k \tilde{f}(\tilde{x}_k, \tilde{p}_k) \tag{$\tilde{x}$-update}$$

We consider generic strategy updates (x-update) and ($\tilde{x}$-update) such that the new strategy profile $f(x_k, p_k) = (f_i(x_k, p_k))_{i\in\mathcal{I}}$ and $\tilde{f}(\tilde{x}_k, \tilde{p}_k) = (\tilde{f}_i(\tilde{x}_k, \tilde{p}_k))_{i\in\tilde{\mathcal{I}}}$ can incorporate a variety of strategy update rules. Two simple examples of such updates include:

1. *Equilibrium update:* The strategy update incorporates a Nash equilibrium strategy profile with respect to the incentive vector in stage k. That is, $f(x_k, p_k) = x^*(p_k)$ and $\tilde{f}(\tilde{x}_k, \tilde{p}_k) = \tilde{x}^*(\tilde{p}_k)$.

2. *Best response update:* The strategy update incorporates a best response strategy with respect to the strategy and the incentive vector in the previous step, i.e. $f_i(x_k, p_k) = \mathrm{BR}_i(x_k, p_k) = \arg\min_{y_i\in X_i} c_i(y_i, x_{-i,k}, p_k)$, $\tilde{f}_i(\tilde{x}_k, \tilde{p}_k) = \tilde{\mathrm{BR}}_i(\tilde{x}_k, \tilde{p}_k) = \arg\min_{\tilde{y}_i\in\tilde{X}_i} \tilde{y}_i^\top \tilde{c}_i(\tilde{x}_k, \tilde{p}_k)$.

Incentive update. In each step $k+1$, the updated incentive vector is a linear combination of the previous vector in step k (i.e. p_k in G and $\tilde{p}_k$ in $\tilde{G}$), and the externality (i.e. $e(x_k)$ in G and $\tilde{e}(\tilde{x}_k)$ in $\tilde{G}$) based on the strategy profile in step k. The relative weight in the linear combination is determined by the step size $\beta_k \in (0,1)$.

$$p_{k+1} = (1-\beta_k)p_k + \beta_k e(x_k); \qquad (p\text{-update})$$

$$\tilde{p}_{k+1} = (1-\beta_k)\tilde{p}_k + \beta_k \tilde{e}(\tilde{x}_k); \qquad (\tilde{p}\text{-update})$$

The incentive updates (p-update)-($\tilde{p}$-update) modify the incentives on the basis of the externality caused by the players. We emphasize that this update is adaptive to the evolution of players' strategies since the externality is evaluated based on players' current strategies. Moreover, the computation of each player's externality only requires that the social planner knows the gradients of its own costs and those of the players, evaluated at the players' current strategy profile. The joint evolution of strategy profiles and incentive vectors $(x_k, p_k)_{k=1}^{\infty}$ (resp. $(\tilde{x}_k, \tilde{p}_k)_{k=1}^{\infty}$) in the atomic game G (resp. non-atomic game $\tilde{G}$) is governed by the learning dynamics (x-update) – (p-update) (resp. ($\tilde{x}$-update) – ($\tilde{p}$-update)). The step-sizes $(\gamma_k)_{k=1}^{\infty}$ and $(\beta_k)_{k=1}^{\infty}$ determine the speed of strategy updates and incentive updates. We make the following assumption on step-sizes:

Assumption 3.1.2.

(i) $\sum_{k=1}^{\infty} \gamma_k = \sum_{k=1}^{\infty} \beta_k = +\infty$, $\sum_{k=1}^{\infty} \gamma_k^2 + \beta_k^2 < +\infty$.

(ii) $\lim_{k\to\infty} \frac{\beta_k}{\gamma_k} = 0$.

Assumption 3.1.2, *(i)* is a standard assumption on step-sizes that allow us to analyze the convergence of the discrete-time learning dynamics. Additionally, *(ii)* assumes that the incentive update occurs at a slower timescale compared to the update of strategies.

Since the assumption on stepsizes (Assumption 3.1.2 *(ii)*) ensures that the incentive evolves on a slower timescale than the strategies, players may view the incentive mechanism as approximately static (although not completely fixed) when updating their strategies. One can show that with any fixed incentive mechanism, strategy updates with Nash equilibrium being the new strategy always converges. On the other hand, although best response updates, which we also consider, do not converge in all games, they converge in many practically-relevant games such as zero sum games [18], potential games [44], and dominance solvable games [32]. Additionally, our strategy updates (x-update) and ($\tilde{x}$-update) can incorporate many other learning dynamics; their convergence properties in static game environments have been extensively studied in the literature, in both atomic and nonatomic games [29], [40], [16], [31].

We emphasize that the convergence of strategy updates with fixed incentive mechanism is not the focus of our paper. Instead, our goal is to characterize conditions under which the adaptive incentive updates (p-update) and ($\tilde{p}$-update) converge to a socially optimal mechanism. We note that such convergence cannot be achieved in scenarios where the strategy updates do not converge even with completely fixed incentive vector. Therefore, we impose the following assumption that the strategy updates consider in our dynamics converge to a Nash equilibrium with any fixed incentive vector.

Assumption 3.1.3. *In G (resp. $\tilde{G}$), the updates* (x-update) *(resp.* ($\tilde{x}$-update)*) starting from any initial strategy x_1 (resp. $\tilde{x}_1$) with $(p_k) \equiv p$ for any p (resp. $(\tilde{p}_k) \equiv \tilde{p}$ for any $\tilde{p}$), satisfies $\lim_{k\to\infty} x_k = x^*(p)$ (resp. $\lim_{k\to\infty} \tilde{x}_k = \tilde{x}^*(\tilde{p})$), where $x^*(p)$ (resp. $\tilde{x}^*(\tilde{p})$) is the Nash equilibrium corresponding to p (resp. $\tilde{p}$).*

3.2 Main Results

In Section 3.2.1 we characterize the set of fixed points of the dynamic updates (x-update)-(p-update) and ($\tilde{x}$-update)-($\tilde{p}$-update), and show that any fixed point corresponds to a socially optimal incentive mechanism such that the induced Nash equilibrium strategy profile minimizes the social cost. In Section 3.2.2, we provide a set of sufficient conditions that guarantee the convergence of strategies and incentives in our learning dynamics. Under these conditions, our learning dynamics designs an adaptive incentive mechanism that eventually induces a socially optimal outcome.

3.2.1 Fixed point analysis

We first characterize the set of fixed points of our learning dynamics (x-update)-(p-update), and ($\tilde{x}$-update)-($\tilde{p}$-update)) as follows:

$$\text{Atomic game } G,\ \{(x,p) | f(x,p) = x,\ e(x) = p\}, \tag{3.10a}$$

$$\text{Nonatomic game } \tilde{G},\ \left\{(\tilde{x},\tilde{p}) | \tilde{f}(\tilde{x},\tilde{p}) = \tilde{x},\ \tilde{e}(\tilde{x}) = \tilde{p}\right\}. \tag{3.10b}$$

We can check that if the learning dynamics start with a fixed point strategy and incentive vector, then the strategies and incentive vectors remain at that fixed point for all time steps. Moreover, under Assumption 3.1.3, we know that for any incentive vector p (resp. $\tilde{p}$), a strategy profile that satisfies $f(x,p) = x$ (resp. $\tilde{f}(\tilde{x},\tilde{p}) = \tilde{x}$) in game G (resp. $\tilde{G}$) must be a Nash equilibrium $x^*(p)$ (resp. $\tilde{x}^*(p)$). Thus, from (3.10a) – (3.10b), we can write the set of incentive vectors at the fixed point as follows:

$$\text{Atomic game } G,\ P^\dagger = \{(p_i^\dagger)_{i\in\mathcal{I}} | e(x^*(p^\dagger)) = p^\dagger\},$$
$$\text{Nonatomic game } \tilde{G},\ \tilde{P}^\dagger = \{(\tilde{p}_i^\dagger)_{i\in\mathcal{I}} | \tilde{e}(\tilde{x}^*(\tilde{p}^\dagger)) = \tilde{p}^\dagger\}.$$

That is, at any fixed point, the incentive of each player is set to be equal to the externality evaluated at their equilibrium strategy profile.

Our next proposition shows that the fixed point set $P^\dagger$ (resp. $\tilde{P}^\dagger$) is non-empty in G (resp. $\tilde{G}$). Moreover, given any fixed point incentive parameter $p^\dagger \in P^\dagger$ and $\tilde{p}^\dagger \in \tilde{P}^\dagger$, the corresponding Nash equilibrium is socially optimal.

Proposition 3.2.1. *In G (resp. $\tilde{G}$), the set $P^\dagger$ (resp. $\tilde{P}^\dagger$) is non-empty. Additionally, any $p^\dagger \in P^\dagger$ (resp. $\tilde{p}^\dagger \in \tilde{P}^\dagger$) is socially optimal in that $x^*(p^\dagger) = x^\dagger$ (resp. $\tilde{x}^*(\tilde{p}^\dagger) = \tilde{x}^\dagger$).*

This result is especially interesting from perspective of implementation because the existence of the optimal incentives implies that for G there exists a *linear* incentive policy (as in (3.1)) which is optimal. Moreover for $\tilde{G}$ there exists a *constant* incentive policy (as in (3.6)) that is optimal.

The proof of Proposition 3.2.1 is based on Brouwer's fixed point theorem. The boundedness of the strategy space allows us to construct convex compact sets which maps to itself under $e(x^*(\cdot))$ (resp. $\tilde{e}(\tilde{x}^*(\cdot))$) in G (resp. in $\tilde{G}$).

Next, we provide sufficient conditions under which the fixed point set $P^\dagger$ and $\tilde{P}^\dagger$ are singleton.

Proposition 3.2.2. *In an atomic game G, the set $P^\dagger$ is singleton if any one of the following conditions holds:*

(i) The equilibrium strategy profile $x^*(p)$ is in the interior of the strategy set X for any p.

(ii) $\langle e(x) - e(x'), x - x' \rangle > 0$ *for all* $x \neq x'$.

In a non-atomic game $\tilde{G}$, $\tilde{P}^\dagger$ *is singleton if the externality function* $\tilde{e}(\cdot)$ *satisfies Assumption 3.1.1 and condition (ii).*

3.2.2 Convergence to optimal incentive mechanism

The next result provides sufficient conditions for strategies and incentives updates (x-update)-(p-update) and ($\tilde{x}$-update)-($\tilde{p}$-update) to converge to social optimality.

Theorem 3.2.3. *Under Assumptions 3.1.2 and 3.1.3, the sequence of strategies and incentives induced by the discrete-time dynamics* (x-update)-(p-update) *in* G *satisfies*

$$\lim_{k \to \infty} (x_k, p_k) = (x^\dagger, p^\dagger) \tag{3.12}$$

if at least one of the following conditions holds:

(C1) If $e_i(x^*(0)) \geq 0$, *then* $\lim_{p \to \infty} e_i(x^*(p)) - p_i < 0$ *for all* $i \in \mathcal{I}$. *If* $e_i(x^*(0)) \leq 0$, *then* $\lim_{p \to -\infty} e_i(x^*(p)) - p_i > 0$ *for all* $i \in \mathcal{I}$.[3] *Moreover,* $\frac{\partial e_i(x^*(p))}{\partial p_j} > 0$ *for all* $p \in \mathbb{R}^n$ *and all* $i \neq j$.

(C2) There exists a continuously differentiable, positive definite and decrescent function [4] $V(p) : \mathbb{R}^n \to \mathbb{R}_+$ *such that* $V(p^\dagger) = 0$ *and* $V(p) > 0$ *for all* $p \neq p^\dagger$. *Moreover:*

$$\nabla V(p)^\top \left(e(x^*(p)) - p\right) < -\omega(\|p - p^\dagger\|) \quad \forall\ p \neq p^\dagger,$$

where $\omega(\cdot)$ *is strictly increasing, and satisfies* $\omega(0) = 0$.

Analogously, the sequence of strategies and incentives in $\tilde{G}$ *induced by* ($\tilde{x}$-update) *and* ($\tilde{p}$-update) *satisfies* $\lim_{k \to \infty}(\tilde{x}_k, \tilde{p}_k) = (\tilde{x}^\dagger, \tilde{p}^\dagger)$ *if the externality function* $\tilde{e}$ *satisfies at least one of (C1) and (C2).*

Owing to Assumption 3.1.2, we utilize the timescale separation between the strategy update (x-update) and the incentive update (p-update) to prove Theorem 3.2.3. Indeed, the two-timescale stochastic approximation theory [4] suggests that the strategy update (x-update) is a fast transient while the incentive update (p-update) is a slow component. Therefore

[3] $p \to \infty$ means $p_i \to \infty$ for all i.

[4] A function $V : \mathbb{R}^n \to \mathbb{R}$ is positive definite if $V(x) \geq \alpha_1(\|x\|)$ for some continuous, strictly increasing function $\alpha_1(\cdot)$ such that $\alpha_1(0) = 0$, and $\alpha_1(t) \to \infty$ as $t \to \infty$. V is decrescent if $V(x) \leq \alpha_2(\|x\|)$ for some continuous, strictly increasing function $\alpha_2(\cdot)$ such that $\alpha_2(0) = 0$.

while considering the fast strategy update one should expect that slow incentive updates are quasi-static. Consequently, Assumption 3.1.3 in game G along with Assumption 3.1.2 ensures that the tuple (x_k, p_k) converges to the set $\{(x^*(p), p) : p \in \mathbb{R}^{|\mathcal{I}|}\}$ [4]. Thus for sufficiently large values of k, the update x_k closely tracks $x^*(p_k)$. Therefore, we consider the following update to analyze the convergence of the slow incentive update (p-update):

$$\mathbf{p}_{k+1} = (1 - \beta_k)\mathbf{p}_k + \beta_k e(x^*(\mathbf{p}_k)). \tag{3.13}$$

Since the step sizes $\{\beta_k\}$ are asymptotically going to zero and have infinite travel (Assumption 3.1.2-(i)) we can approximate the updates in (B.24) by the following continuous-time dynamical system:

$$\dot{\mathbf{p}}(t) = e(x^*(\mathbf{p}(t))) - \mathbf{p}(t), \tag{3.14}$$

The convergence of discrete-time updates (x-update)-(p-update) then hold if the flow of (3.14) globally converges to $P^\dagger$.

Requirements (C1) in Theorem 3.2.3 are sufficient conditions for convergence of the trajectories of (3.14) to the set $P^\dagger$. This condition is based on cooperative dynamical systems theory [19]. Intuitively, condition (C1) demands that in the equilibrium if the players inflict (resp. alleviate) some externality when no incentive is applied then there should exist high enough prices (resp. subsidies) which can compensate for the externality. Moreover, it demands that higher prices (resp. subsidies) on other player increases the externality inflicted (resp. alleviated) by a player.

Requirement (C2) in Theorem 3.2.3 on the other hand ensures convergence by positing existence of a Lyapunov function [41] that is strictly positive everywhere except at $P^\dagger$ and decreases along the flow of (3.14).

Note that either one of the conditions (C1) or (C2) guarantees the convergence of the flow of the slow system (3.14) to $P^\dagger$. This in addition to the convergence of the fast strategy update (Assumption 3.1.3) leads to the convergence of the discrete-time dynamics (x-update)-(p-update) [4, Chapter 6].

Thus, we have shown that there exists an incentive which induces an equilibrium which is socially optimal and the externality based pricing update along with any strategy update, satisfying requirements of Theorem 3.2.3, converges to the optimal incentive.

3.3 Applications

In this section, we apply our general results to three classes of games that are practically relevant: (Section 3.3.1) Atomic networked aggregative games; (Section 3.3.2) Atomic Cournot games; and (Section 3.3.3) Non-atomic routing games. In each case, we show that our adaptive incentive mechanism asymptotically induces a socially optimal outcome.

3.3.1 Atomic networked aggregative games

We consider a finite set of players $\mathcal{I}$ who are connected in a network. The strategy of each player $i \in \mathcal{I}$ is a real number $x_i \in \mathbb{R}$. We represent the network that connects players by a matrix $w = (w_{ij})_{i,j \in \mathcal{I}}$, where w_{ij} captures the impact of player j's strategy x_j on player i's cost. We assume that $w_{ii} = 0$ for all $i \in \mathcal{I}$. The cost of each player $i \in \mathcal{I}$ given any strategy profile $x = (x_i)_{i \in \mathcal{I}}$ is a quadratic function as follows:

$$\ell_i(x) = \frac{1}{2}x_i^2 - a_i x_i z_i(x)$$

where $a_i > 0$ and $z_i(x) = \sum_{j \in \mathcal{I}} w_{ij} x_j$ is the average strategy of player i's neighbors weighted by the network matrix w. That is, $z_i(x)$ captures the network effect of opponents' strategies on player i.

Networked aggregative games are applicable in a variety of settings, where players' strategies and costs are affected by those around them. Examples of such settings include peer effects, investment in networked markets, and cross-neighborhood impacts of crime [21].

A social planner designs an incentive mechanism that charges each player i with payment $p_i x_i$. The total cost of player i under strategy profile x and p is

$$c_i(x, p) = \frac{1}{2}x_i^2 - x_i a_i z_i(x) + p_i x_i.$$

Let $A = \mathsf{diag}([a_1, a_2, ..., a_{|\mathcal{I}|}]^\top)$. We assume that the matrix $(I - Aw)$ is invertible and let $L = (I - Aw)^{-1}$. In the economics literature, the matrix $(I - Aw)^{-1}$ is referred to as the *Leontief matrix*, where the ij entry of this matrix captures how the payment of player j affects the equilibrium strategy of player i ([34]).

For any p, we show that the aggregative game has a unique Nash equilibrium given by:

$$x^*(p) = -(I - Aw)^{-1}p. \tag{3.15}$$

Given x, the cost of the social planner is $\Phi(x) = \frac{1}{2}\sum_{i\in I}(x_i - \xi_i)^2$ for $\xi_i \in \mathbb{R}$, where $x^\dagger = (\xi_i)_{i\in I}$ is the planner's socially optimal strategy profile. Moreover, from (3.5), the externality caused by player i given the strategy profile x is $e_i(x) = \xi_i + a_i z_i(x)$.

We consider the learning dynamics, where players update their strategies using best response. Given any strategy profile x and incentive p, the best response of player i is $BR_i(x_{-i}, p_i) = a_i z_i(x) - p_i$ and the best response vector $BR(x, p) = (BR_i(x_{-i}, p_i))_{i\in\mathcal{I}}$. Thus the strategy update is $f(x, p) = BR(x, p) = Awx - p$. Then, the discrete-time leaning dynamics (x-update) – (p-update) can be written as follows:

$$x_{k+1} = (1 - \gamma_k)x_k + \gamma_k\left(Awx_k - p_k\right), \tag{3.16a}$$

$$p_{k+1} = (1 - \beta_k)p_k + \beta_k\left(\xi + Awx_k\right), \tag{3.16b}$$

and the step-sizes $\{\gamma_k\}_{k=1}^\infty, \{\beta_k\}_{k=1}^\infty$ satisfy Assumption 3.1.2.

We show that there exists a unique $p^\dagger$ such that the induced equilibrium strategy profile $x^*(p^\dagger)$ equals to the socially optimal strategy $x^\dagger$. Moreover, we also provide sufficient condition on the Leontief matrix under which our learning dynamics converges the unique socially optimal incentive mechanism.

Proposition 3.3.1. *The unique socially optimal incentive mechanism is* $p^\dagger = (I - Aw)\xi$. *Furthermore if the real part of eigenvalues of* $(I - Aw)$ *is positive, i.e.* $\mathsf{spec}(L) \subset \mathbb{C}_+^\circ$, *then the discrete-time learning dynamics* (3.16a)-(3.16b) *satisfy* $\lim_{k\to\infty}(x_k, p_k) = (x^\dagger, p^\dagger)$.

From (3.15), we know that for any p, there exists a unique equilibrium strategy profile $x^*(p)$ that is linear in p. Then, we obtain the socially optimal incentive $p^\dagger$ that satisfies $x^*(p^\dagger) = x^\dagger$.

Additionally, $p^\dagger$ also satisfies that $e(x^\dagger(p^\dagger)) = p^\dagger$, and therefore $p^\dagger$ is a fixed point of the learning dynamics. To show the convergence results in Proposition 3.3.1 we verify the conditions in Theorem 3.2.3 holds. The condition $\mathsf{spec}(L) \subset \mathbb{C}_+^\circ$ ensures that Assumption 3.1.3 holds. Indeed, we show that the strategy update (3.16a) with fixed incentives asymptotically track the flow of a continuous-time linear dynamical system. The condition $\mathsf{spec}(L) \subset \mathbb{C}_+^\circ$ ensures that the flow of continuous-time dynamical system asymptotically converges to fixed points of (3.16a) with fixed incentives. Finally, we verify that if $\mathsf{spec}(L) \subset \mathbb{C}_+^\circ$ then $K = -L \subset \mathbb{C}_-^\circ$ and there exists a Lyapunov function that satisfies (C2) in Theorem 3.2.3. In particular, the Lyapunov function is given by

$$V(p) = (p - p^\dagger)^\top M (p - p^\dagger)$$

where M is a matrix that satisfies $K^\top M + MK = -I$.[5]

3.3.2 Atomic Cournot competition

A finite set of firms $\mathcal{I}$ compete in a single market. The strategy of each firm $i \in \mathcal{I}$ is its production quantity x_i. Given any strategy profile $x = (x_i)_{i\in\mathcal{I}}$, the price of the good is $\xi(x) = \theta - \delta \sum_{i\in\mathcal{I}} x_i$ with $\theta, \delta > 0$. The per-unit production cost of the good is ν. Then, the cost function of firm $i \in \mathcal{I}$ (written as negative of the profit) is given by:

$$\ell_i(x) = -x_i \xi(x) + \nu x_i \tag{3.17}$$

A social planner designs an incentive mechanism that charges each player i with payment $p_i x_i$. The total cost of firm $i \in \mathcal{I}$ given x and p is:

$$c_i(x, p) = -x_i \xi(x) + (\nu + p_i) x_i$$

The game has a unique Nash equilibrium given by: [6]

$$x_i^*(p) = \frac{1}{\delta(|\mathcal{I}| + 1)} \left(\theta - \nu - |\mathcal{I}| p_i + \sum_{j \neq i} p_j \right) \tag{3.18}$$

The goal of the social planner is to minimize the aggregate cost of players while also accounting for the environmental cost of good production, which is unpriced in equilibrium. We model the environmental cost to be a quadratic function of production following [7].

[5] The existence of a matrix M is guaranteed by the Lyapunov theorem [8] as $\mathsf{spec}(K) \subset \mathbb{C}_-^\circ$.

[6] We assume that θ is large enough such that $x^*(p) > 0$ for all p in a neighborhood of the socially optimal incentive $p^\dagger$.

Thus, the social cost function is $\Phi(x) = \sum_{i=1}^{n} \ell_i(x) + \lambda \sum_{i=1}^{n} x_i^2$ where $\lambda > 0$ is a parameter that determines the relative weight between the firm costs and environmental cost. Finally, the externality (3.5) caused by of a firm $i \in \mathcal{I}$ is $e_i(x) = 2\lambda x_i + \delta \sum_{j \neq i} x_j$.

We consider best response strategy updates. Given any x_{-i}, the best response of firm $i \in \mathcal{I}$ is:

$$BR_i(x_{-i}, p_i) = \frac{\theta - \delta \sum_{j \neq i} x_j - \nu - p_i}{2\delta}.$$

Following (x-update) – (p-update), we can write the updates of strategies and incentives as follows:

$$x_{i,k+1} = (1-\gamma_k)x_{i,k} + \gamma_k \left(\frac{\theta - \delta \sum_{j \neq i} x_{j,k} - \nu - p_{i,k}}{2\delta} \right), \tag{3.19a}$$

$$p_{i,k+1} = (1-\beta_k)p_{i,k} + \beta_k(\delta \sum_{j \neq i} x_{j,k} + 2\lambda x_{i,k}). \tag{3.19b}$$

and the step-sizes $\{\gamma_k\}_{k=1}^{\infty}, \{\beta_k\}_{k=1}^{\infty}$ satisfy Assumption 3.1.2.

We can show that for any fixed p, the best response learning dynamics (3.19a) converges to a Nash equilibrium $x^*(p)$ associated with p. Indeed, we show that the strategy update (3.19a) with fixed incentives asymptotically track the flow of a continuous-time linear dynamical system whose flow asymptotically converges to the Nash equilibrium $x^*(p)$. Thus, Assumption 3.1.3 is satisfied.

The next proposition shows that the optimal incentive $p^\dagger$ is unique. Moreover, the incentive vectors induced by (3.19b) converge to the socially optimal incentive $p^\dagger$ if the weight of environmental cost, λ, is sufficiently high.

Proposition 3.3.2. *There exists a unique socially optimal incentive mechanism $p^\dagger$ that satisfies $p^\dagger = e(x^*(p^\dagger))$. Given $p^\dagger$, the induced equilibrium strategy profile is socially optimal, i.e. $x^*(p^\dagger) = x^\dagger$. Moreover, the discrete-time learning dynamics* (3.19a)-(3.19b) *satisfy* $\lim_{k\to\infty}(x_k, p_k) = (x^\dagger, p^\dagger)$ *if $\lambda > \delta$.*

Recall that λ is the weight of environmental cost in the social cost function, and δ is the increase of firm cost with respect to the increase of production level. The sufficient condition $\lambda > \delta$ states that if the social planner assigns higher weight to the environmental cost compared to the per-unit increase of firm cost, then the adaptive incentive mechanism can asymptotically induce a socially optimal outcome.

The proof of Proposition 3.3.2 follows similarly to that of Proposition 3.3.1. We show that there is a unique incentive $p^\dagger$ such that the corresponding Nash equilibrium as in (3.19a) equals to the socially optimal strategy profile, and $p^\dagger$ is a fixed point of the discrete-time learning dynamics (3.19b). Moreover, we show that when $\lambda > \delta$, we can construct a Lyapnov function that satisfies (C2) in Assumption (3.1.3). Therefore, following Theorem (3.2.3), we can conclude that the discrete-time learning dynamics converges to a socially optimal outcome.

3.3.3 Non-atomic routing games

A traveler population with total demand of 1 make routing decisions on a parallel-route network, where a single origin - destination pair is connected by a finite set of routes S. The strategy of the traveler population is $\tilde{x} = (\tilde{x}^j)_{j \in S}$, where $\tilde{x}^j$ is the mass of travelers who choose route $j \in S$. The population's strategy set is $\tilde{X} = \{\tilde{x} | \sum_{j \in S} \tilde{x}^j = 1, \ \tilde{x}^j \geq 0, \ \forall j \in S\}$.

Given any $\tilde{x}$ and any route $j \in S$, the travel time cost $\ell^j(\tilde{x}^j)$ is a *strictly-increasing* and *convex* function of the mass of travelers who take route $j \in S$. This reflects the congestible nature of the traffic routes and the fact that the travel time increases faster as more travelers take that route.

A social planner designs a tolling mechanism $\tilde{p} = (\tilde{p}^j)_{j \in S}$, where the toll price of route j is $\tilde{p}^j$. Given any $\tilde{x}$ and $\tilde{p}$, the total cost experienced by travelers who take route j is $\tilde{c}^j(\tilde{x}, \tilde{p}) = \ell^j(\tilde{x}^j) + \tilde{p}^j$.

Given any toll vector $\tilde{p}$, the routing game has a unique Nash equilibrium $\tilde{x}^*(\tilde{p})$. The goal of the social planner is to minimize the total cost of all routes in the network, i.e. $\tilde{\Phi}(\tilde{x}) = \sum_{j \in S} \tilde{x}^j \ell_j(\tilde{x}^j)$. We can check that $\tilde{\Phi}(\tilde{x})$ is strictly convex in $\tilde{x}$, and thus the socially optimal strategy $\tilde{x}^\dagger$ is unique. Finally, following from (3.9), the externality caused by travelers on route $j \in S$ is $\tilde{e}^j(\tilde{x}) = \tilde{x}^j \frac{d\ell^j(\tilde{x}^j)}{d\tilde{x}^j}$.

We consider perturbed best response strategy updates. Given any $\tilde{x}$, the perturbed best response strategy is

$$\tilde{f}^i(\tilde{x}, \tilde{p}) = \frac{\exp(-\eta \tilde{c}^i(\tilde{x}, \tilde{p}))}{\sum_{j \in S} \exp(-\eta \tilde{c}^j(\tilde{x}, \tilde{p}))},$$

where η evaluates the sensitivity of travelers' route choices with respect to the costs. We note that as $\eta \to \infty$, the perturbed best response strategy reduces to a best response strategy that only chooses routes with the minimal costs.

The discrete-time learning dynamics are:

$$\tilde{x}^j_{k+1} = (1 - \gamma_k)\tilde{x}^j_k + \gamma_k \frac{\exp(-\eta \tilde{c}^j(\tilde{x}_k, \tilde{p}_k))}{\sum_{j \in S} \exp(-\eta \tilde{c}^j(\tilde{x}_k, \tilde{p}_k))}, \tag{3.20a}$$

$$\tilde{p}^j_{k+1} = (1 - \beta_k)\tilde{p}^j_k + \beta_k \tilde{x}^j_k \frac{d\ell^j(\tilde{x}^j_k)}{d\tilde{x}^j_k}. \tag{3.20b}$$

and the step-sizes $\{\gamma_k\}_{k=1}^{\infty}, \{\beta_k\}_{k=1}^{\infty}$ satisfy Assumption 3.1.2. Moreover, we can show that for any fixed $\tilde{p}$, the perturbed best response dynamics (3.20a) converges to the perturbed equilibrium. Indeed, due to Assumption 3.1.2-i) the discrete-time updates (3.20a), with fixed incentive $\tilde{p}$, tracks the flow of a cooperative continuous-time dynamical system [19] whose flows converges to the perturbed equilibrium. Thus Assumption 3.1.3 holds.

The next proposition shows that the optimal incentive $\tilde{p}^\dagger$ is unique. Moreover, the incentive vectors induced by (3.20b) converge to the socially optimal incentive $p^\dagger$.

Proposition 3.3.3. *As $\eta \to \infty$, the strategies and incentives induced by the discrete-time learning dynamics* (3.20a)-(3.20b) *converge to a unique fixed point, i.e.* $\lim_{k\to\infty}(\tilde{x}_k, \tilde{p}_k) = (\tilde{x}^*(\tilde{p}^\dagger), \tilde{p}^\dagger)$*, where $\tilde{x}^*(p^\dagger)$ is a Nash equilibrium given $p^\dagger$, and $p^\dagger$ satisfies $\tilde{e}(\tilde{x}^*(\tilde{p}^\dagger)) = \tilde{p}^\dagger$. Additionally, $\tilde{p}^\dagger$ is the unique optimal incentive mechanism, and the corresponding Nash equilibrium $\tilde{x}^*(p^\dagger) = \tilde{x}^\dagger$.*

In the proof of Proposition 3.3.3, we first show that the externality function $\tilde{e}$ is monotonic in $\tilde{x}$. Thus, the existence and uniqueness of fixed point toll price follows from Propositions 3.2.1 and 3.2.2. Additionally, we show that the value of externality in equilibrium $\tilde{e}(\tilde{x}^*(p))$ satisfies (C1) in Theorem 3.2.3. Therefore, we can conclude that the discrete-time learning dynamics converge to the socially optimal outcome.

Chapter 4

Conclusion

In this thesis, we proposed tolling and incentive updates that can allow societal-scale system designers to steer selfish players to social optimality in situations where the players are dynamically responding to changing system conditions and other players' strategies. In particular, we demonstrated the following:

- In Chapter 2, we proposed a two timescale discrete-time stochastic dynamics that captures the joint evolution of loads in parallel networks and adaptive adjustment of tolls [28]. We analyzed the properties of fixed points and the convergence of loads and tolls in this dynamics. Under our dynamics, the tolls asymptotically induce a socially optimal load with high probability. Our results allow a central authority to set tolls in traffic networks *adaptively* and *optimally* in response to the dynamic arrival of travelers who myopically and selfishly make routing decisions.

- In Chapter 3, we proposed a joint strategy and incentive update scheme for general atomic and nonatomic games so that the emergent Nash equilibrium minimizes a social planner's cost (or equivalently maximizes social welfare) [27]. We assume that the planner, at each time-step, can modify the costs of players by setting a payment. There are three key features of the proposed scheme: first, the incentives are updated at a slower timescale as compared to the players' strategy update. Second, the incentive update is based on the externality caused by the players' strategy evaluated as the difference between players' marginal cost and the planner's marginal cost. Third, the incentive update is agnostic to the specific strategy update deployed by players, and relies on the current strategy profile. We showed applications to atomic networked quadratic aggregative games, atomic Cournot competition and nonatomic routing games.

Combined, these provide a general scheme for system designers to adaptively incentivize selfish players and steer them to social optimality over time. Both the tolling scheme for traffic networks and incentives for general games focus on updates that have a natural economic meaning as the externality caused by players on others. Further, the updates can be implemented using available information in both cases (the loads on the edges of the traffic network and the strategies of the players, respectively).

A direction for future work in the case of traffic routing is to extend the results from parallel networks to general networks, and account for the fact that travelers may change their route choices at intermediate nodes in the network. In the case of incentives in general games, it remains to be understood for what class of games the sufficient conditions for convergence of the incentive updates are satisfied.

Bibliography

[1] Jorge Barrera and Alfredo Garcia. “Dynamic incentives for congestion control”. In: *IEEE Transactions on Automatic Control* 60.2 (2014), pp. 299–310.

[2] Tamer Başar. “Affine incentive schemes for stochastic systems with dynamic information”. In: *SIAM Journal on Control and Optimization* 22.2 (1984), pp. 199–210.

[3] Vivek S Borkar. “Cooperative dynamics and Wardrop equilibria”. In: *Systems & control letters* 58.2 (2009), pp. 91–93.

[4] Vivek S Borkar. “Stochastic approximation with two time scales”. In: *Systems & Control Letters* 29.5 (1997), pp. 291–294.

[5] Vivek S Borkar. *Stochastic approximation: a dynamical systems viewpoint*. Vol. 48. Springer, 2009.

[6] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[7] Desmond Cai, Subhonmesh Bose, and Adam Wierman. “On the role of a market maker in networked cournot competition”. In: *Mathematics of Operations Research* 44.3 (2019), pp. 1122–1144.

[8] Frank M Callier and Charles A Desoer. *Linear system theory*. Springer Science & Business Media, 2012.

[9] George Christodoulou and Elias Koutsoupias. “The price of anarchy of finite congestion games”. In: *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*. 2005, pp. 67–73.

[10] Roberto Cominetti, Francisco Facchinei, and Jean B Lasserre. *Modern Optimization Modelling Techniques*. Springer Science & Business Media, 2012.

[11] Roberto Cominetti, Emerson Melo, and Sylvain Sorin. “A payoff-based learning procedure and its application to traffic games”. In: *Games and Economic Behavior* 70.1 (2010), pp. 71–83.

[12] Giacomo Como and Rosario Maggistro. “Distributed Dynamic Pricing of Multiscale Transportation Networks”. In: *IEEE Transactions on Automatic Control* (2021).

[13] Otto A Davis and Andrew Whinston. “Externalities, welfare, and the theory of games”. In: *Journal of Political Economy* 70.3 (1962), pp. 241–262.

[14] Francisco Facchinei and Jong-Shi Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.

[15] Farhad Farokhi and Karl H Johansson. "A piecewise-constant congestion taxing policy for repeated routing games". In: *Transportation Research Part B: Methodological* 78 (2015), pp. 123–143.

[16] Drew Fudenberg and David Levine. *The theory of learning in games*. Vol. 2. MIT press, 1998.

[17] J. E. Gentle. *Matrix Algebra*. Vol. 10. Springer, 2007, pp. 978–.

[18] Christopher Harris. "On the rate of convergence of continuous-time fictitious play". In: *Games and Economic Behavior* 22.2 (1998), pp. 238–259.

[19] Morris W Hirsch. "Systems of differential equations that are competitive or cooperative II: Convergence almost everywhere". In: *SIAM Journal on Mathematical Analysis* 16.3 (1985), pp. 423–439.

[20] Yu-Chi Ho, Peter B Luh, and Geert Jan Olsder. "A control-theoretic view on incentives". In: *Automatica* 18.2 (1982), pp. 167–179.

[21] Matthew O Jackson and Yves Zenou. "Games on networks". In: *Handbook of game theory with economic applications*. Vol. 4. Elsevier, 2015, pp. 95–163.

[22] Soumya Karlamangla. "Why Evening Rush Hour Feels So Much Worse Now". In: *New York Times* (2021). Published on August 11, 2021. Available at `https://www.nytimes.com/2021/08/11/us/ca-rush-hour-traffic.html`.

[23] Robert Kleinberg, Georgios Piliouras, and Éva Tardos. "Multiplicative updates outperform generic no-regret learning in congestion games". In: *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 2009, pp. 533–542.

[24] Walid Krichene, Benjamin Drighes, and Alexandre Bayen. "On the convergence of no-regret learning in selfish routing". In: *International Conference on Machine Learning*. PMLR. 2014, pp. 163–171.

[25] Dan Levin. "Taxation within Cournot oligopoly". In: *Journal of Public Economics* 27.3 (1985), pp. 281–290.

[26] Boyi Liu, Jiayang Li, Zhuoran Yang, Hoi-To Wai, Mingyi Hong, Yu Marco Nie, and Zhaoran Wang. "Inducing Equilibria via Incentives: Simultaneous Design-and-Play Finds Global Optima". In: *arXiv preprint arXiv:2110.01212* (2021).

[27] Chinmay Maheshwari, Kshitij Kulkarni, Manxi Wu, and Shankar Sastry. "Inducing Social Optimality in Games via Adaptive Incentive Design". In: *arXiv preprint arXiv:2204.05507* (2022).

[28] Chinmay Maheshwari*, Kshitij Kulkarni*, Manxi Wu, and Shankar Sastry. "Dynamic Tolling for Inducing Socially Optimal Traffic Loads". In: *arXiv preprint arXiv:2110.08879* (2021).

[29] Eric Mazumdar, Lillian J Ratliff, and S Shankar Sastry. "On gradient-based learning in continuous games". In: *SIAM Journal on Mathematics of Data Science* 2.1 (2020), pp. 103–131.

[30] Emily Meigs, Francesca Parise, and Asuman Ozdaglar. "Learning dynamics in stochastic routing games". In: *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2017, pp. 259–266.

[31] Dov Monderer and Lloyd S Shapley. "Fictitious play property for games with identical interests". In: *Journal of economic theory* 68.1 (1996), pp. 258–265.

[32] John H Nachbar. ""Evolutionary" selection dynamics in games: Convergence and limit properties". In: *International journal of game theory* 19.1 (1990), pp. 59–89.

[33] Dario Paccagnan, Rahul Chandan, Bryce L Ferguson, and Jason R Marden. "Incentivizing efficient use of shared infrastructure: Optimal tolls in congestion games". In: *arXiv preprint arXiv:1911.09806* (2019).

[34] Francesca Parise and Asuman Ozdaglar. "Sensitivity analysis for network aggregative games". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2017, pp. 3200–3205.

[35] Arthur Cecil Pigou and Nahid Aslanbeigui. *The economics of welfare*. Routledge, 2017.

[36] Jorge I Poveda, Philip N Brown, Jason R Marden, and Andrew R Teel. "A class of distributed adaptive pricing mechanisms for societal systems with limited information". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2017, pp. 1490–1495.

[37] Lillian J Ratliff and Tanner Fiez. "Adaptive incentive design". In: *IEEE Transactions on Automatic Control* 66.8 (2020), pp. 3871–3878.

[38] Tim Roughgarden. "Algorithmic game theory". In: *Communications of the ACM* 53.7 (2010), pp. 78–86.

[39] Tim Roughgarden and Éva Tardos. "How bad is selfish routing?" In: *Journal of the ACM (JACM)* 49.2 (2002), pp. 236–259.

[40] William H Sandholm. *Population games and evolutionary dynamics*. MIT press, 2010.

[41] Shankar Sastry. *Nonlinear systems: analysis, stability, and control*. Vol. 10. Springer Science & Business Media, 2013.

[42] Rachel Schnalzer. "Traffic is terrible again. Here's how to get it closer to spring 2020 levels." In: *LA Times* (2021). Published on July 22, 2021. Available at `https://www.latimes.com/business/story/2021-07-22/los-angeles-traffic-congestion-commute-pandemic`. (Visited on 08/20/2021).

[43] Guni Sharon, Josiah P Hanna, Tarun Rambha, Michael W Levin, Michael Albert, Stephen D Boyles, and Peter Stone. "Real-time adaptive tolling scheme for optimized social welfare in traffic networks". In: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2017)*. 2017.

[44] Brian Swenson, Ryan Murray, and Soummya Kar. "On best-response dynamics in potential games". In: *SIAM Journal on Control and Optimization* 56.4 (2018), pp. 2734–2767.

[45] Richard S Varga. *Geršgorin and his circles*. Vol. 36. Springer Science & Business Media, 2010.

[46] Hal R Varian. "A solution to the problem of externalities when agents are well-informed". In: *The American Economic Review* (1994), pp. 1278–1293.

[47] Manxi Wu, Saurabh Amin, and Asuman Ozdaglar. "Bayesian Learning with Adaptive Load Allocation Strategies". In: *Learning for Dynamics and Control*. PMLR. 2020, pp. 561–570.

Appendix A

Convergence for cooperative dynamical systems

In this section we review a result from [19] which provides an easily verifiable requirement on the vector field which ensures convergence to equilibrium.

Definition A.0.1. *A dynamical system*

$$\dot{x}(t) = f(x(t)) \tag{A.1}$$

with a C^1- *function* $f : \mathbb{R}^n \to \mathbb{R}^n$ *is* cooperative *if:*

(P-i) *for any* $i \neq j \in [n]$ *and* $x \in \mathbb{R}^n$ *we have* $\frac{\partial f_i(x)}{\partial x_j} \geq 0$*;*

(P-ii) *the Jacobian matrix* $\nabla f(x)$ *is irreducible;*[1]

(P-iii) *for every* $i \in [n]$, $f_i(0) \geq 0$*;*

(P-iv) *for any* $x \in \mathbb{R}^n_{\geq 0}$ *there exists* $y > x$ *with* $f_i(y) < 0$ *for all* $i \in [n]$

Theorem A.0.2 ([19, Theorem 5.1])**.** *If the dynamical system* (A.1) *is cooperative (i.e. satisfies the assumptions (P-i)-(P-iv)), trajectories starting almost everywhere in* $\mathbb{R}^n_{\geq 0}$ *converge to the set* $\{x : f(x) = 0\}$*.*

[1] A matrix $A = [A_{ij}] \in \mathbb{R}^{n \times n}$ is irreducible if whenever the set $\{1, \ldots, n\}$ is expressed as the union of two disjoint proper subsets S, S', then for every $i \in S$ there exist $j, k \in S'$ such that $A_{ij} \neq 0$ and $A_{ki} \neq 0$.

Appendix B

Proofs

B.0.1 Proof of Lemma 2.2.4

In the subsequent proof we shall show that the following claims hold:

(C-I) For any $p \in \mathbb{R}^R$, the optimizer of (2.11) is unique; call it $y^\star(p)$;

(C-II) $y^\star(p)$ is an optimal solution if and only if it satisfies $h(y^\star(p), p) = y^\star$

Note that the above two claims ensure that $\bar{x}^{(\beta)}(p) = y^\star(p)$. The subsequent exposition establishes the validity of these claims.

To see that (C-I) is true, we note that the feasible set of (2.11) is a compact convex set. Moreover, from the convexity of ℓ it follows that $\sum_{i\in[R]} \int_0^{y_i} c_i(s, p_i)ds + \frac{1}{\beta} \sum_{i\in[R]} y_i \ln y_i$ is a strictly convex function in y. Strict convexity ensures the uniqueness of solution of a convex optimization problem on a convex compact set [6].

For (C-II), we now employ KKT conditions of optimality: let $\delta \in \mathbb{R}$ be the Lagrange multiplier corresponding to the equality constraint. Define the Lagrangian as follows:

$$L(y, \delta, p) \coloneqq \sum_{i\in[R]} \int_0^{y_i} c_i(s, p_i)ds + \frac{1}{\beta} \sum_{i\in R} y_i \ln y_i + \delta \left(\sum_{i\in[R]} y_i - \frac{\lambda}{\mu} \right)$$

The optimal solution $y^\star(p)$ and Lagrange multiplier $\delta^\star$ satisfy:

1. $\nabla_y L(y^\star(p), \delta^\star, p) = 0$. This gives us $c_i(y_i^\star(p), p_i) + \frac{1}{\beta}(1 + \ln(y_i^\star(p))) + \delta^\star = 0$, which can

be also written as follows

$$y_i^\star(p) = \exp(-\beta\delta^\star - 1)\exp(-\beta c_i(y_i^\star(p), p_i)), \quad \forall\ i \in [R]. \tag{B.1}$$

2. $\nabla_\delta L(y^\star(p), \delta^\star, p) = 0$. That is, $\sum_{i\in R} y_i^\star(p) = \frac{\lambda}{\mu}$.

From (B.1), summing over $i \in [R]$ on both the sides, we obtain

$$\frac{\lambda}{\mu} = \exp(-\beta\delta^\star - 1)\sum_{i\in[R]} \exp(-\beta c_i(y_i^\star(p), p_i))$$

The above conditions give us that $y^\star(p)$ is an optimal solution of (2.11) if and only if it satisfies:

$$y_i^\star(p) = \frac{\lambda}{\mu}\frac{\exp(-\beta c_i(y_i^\star(p), p_i))}{\sum_{i\in[R]}\exp(-\beta c_i(y_i^\star(p), p_i))} = h_i(y^\star(p), p), \quad \forall\ i \in [R].$$

B.0.2 Proof of Lemma 2.2.5

Before proving Lemma 2.2.5, we will first state the following result that will be useful later. Define $\mathbf{C}(x) \coloneqq \mathsf{diag}\left(\left(\frac{d\ell_i(x)}{dx}\right)_{i=1}^R\right)$ and $\mathbf{M}(x) \coloneqq \mathsf{diag}\left(\left(x_i\frac{d\ell_i(x)}{dx}\right)_{i=1}^R\right)$.

Lemma B.0.1. *The function $h(x, p)$ presented in* (2.5) *is Lipschitz in x and p and satisfies:*

$$\nabla_x h(\bar{x}^{(\beta)}(p), p) = \frac{\lambda}{\mu}\beta\bar{x}^{(\beta)}(p)\bar{x}^{(\beta)}(p)^\top \boldsymbol{C}(\bar{x}^{(\beta)}(p)) - \beta\boldsymbol{M}\left(\bar{x}^{(\beta)}(p)\right) \tag{B.2}$$

Proof of Lemma B.0.1. To show that the function is Lipschitz, it is sufficient to show that the norm of the gradient of this function is bounded. This is due to the first order Taylor expansion of the function. Therefore, for every $i \in [R]$ we first compute $\nabla_x h_i(x, p)$ entrywise as follows:

$$\frac{\partial h_i(x,p)}{\partial x_j} = \begin{cases} \frac{\lambda}{\mu}\frac{-\beta\frac{d\ell_i(x_i)}{dx}\sum_{k\in[R]}\exp(-\beta c_k(x_k,p_k))\exp(-\beta c_i(x_i,p_i)) + \beta\frac{d\ell_i(x_i)}{dx}(\exp(-\beta c_i(x_i,p_i)))^2}{\left(\sum_{k\in[R]}\exp(-\beta c_k(x_k,p_k))\right)^2} & \text{if } \ i = j; \\ \frac{\lambda}{\mu}\beta\frac{d\ell_j(x_j)}{dx}\exp(-\beta c_i(x_i,p_i))\frac{\exp(-\beta c_j(x_j,p_j))}{\left(\sum_{k\in[R]}\exp(-\beta c_k(x_k,p_k))\right)^2} & \text{otherwise} \end{cases} \tag{B.3}$$

Evaluating the above derivative at $\bar{x}^{(\beta)}(p)$

$$\frac{\partial h_i(\bar{x}^{(\beta)}(p), p)}{\partial x_j} = \begin{cases} \beta\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx}\left(-\bar{x}_i^{(\beta)}(p) + \frac{\mu}{\lambda}\left(\bar{x}_i^{(\beta)}(p)\right)^2\right) & \text{if } \ i = j; \\ \frac{\mu}{\lambda}\beta\frac{d\ell_j(\bar{x}_j^{(\beta)}(p))}{dx}\bar{x}_i^{(\beta)}(p)\bar{x}_j^{(\beta)}(p) & \text{otherwise} \end{cases} \tag{B.4}$$

To state it concisely:

$$\nabla_x h(\bar{x}^{(\beta)}(p), p) = \frac{\lambda}{\mu}\beta\bar{x}^{(\beta)}(p)\bar{x}^{(\beta)}(p)^\top\mathbf{C}(\bar{x}^{(\beta)}(p)) - \beta\mathbf{M}\left(\bar{x}^{(\beta)}(p)\right) \tag{B.5}$$

Similarly,

$$\frac{\partial h_i(x,p)}{\partial p_j} = \begin{cases} \frac{\lambda}{\mu} \frac{-\beta \sum_{k\in[R]} \exp(-\beta c_k(x_k,p_k)) \exp(-\beta c_i(x_i,p_i)) + \beta(\exp(-\beta c_i(x_i,p_i)))^2}{\left(\sum_{k\in[R]} \exp(-\beta c_k(x_k,p_k))\right)^2} & \text{if } \ i=j; \\ \frac{\lambda}{\mu}\beta \exp(-\beta c_i(x_i,p_i)) \frac{\exp(-\beta c_j(x_j,p_j))}{\left(\sum_{k\in[R]} \exp(-\beta c_k(x_k,p_k))\right)^2} & \text{otherwise} \end{cases} \tag{B.6}$$

To conclude the proof we observe that on a bounded domain the derivative in (B.3) and (B.6) are bounded. □

Proof of Lemma 2.2.5

Recall from the proof of Lemma 2.2.4 ((C-II) to be precise) that for any $p \in \mathbb{R}^R$, $\bar{x}^{(\beta)}(p)$ is a solution to the optimization problem (2.11). Let the feasible set in the optimization problem (2.11) be denoted by F. Using first order conditions for constrained optimization we obtain the variational inequality:

$$\sum_{i\in[R]} \left(c_i\left(\bar{x}_i^{(\beta)}(p), p\right) + \frac{1}{\beta}\left(\ln\left(\bar{x}_i^{(\beta)}(p)\right) + 1\right)\right)\left(x_i - \bar{x}_i^{(\beta)}(p)\right) \geq 0 \quad \forall x \in F. \tag{B.7}$$

Similarly writing the above condition for p', we obtain

$$\sum_{i\in[R]} \left(c_i\left(\bar{x}_i^{(\beta)}(p'), p'\right) + \frac{1}{\beta}\left(\ln\left(\bar{x}_i^{(\beta)}(p')\right) + 1\right)\right)\left(y_i - \bar{x}_i^{(\beta)}(p')\right) \geq 0 \quad \forall y \in F. \tag{B.8}$$

Choosing $x = \bar{x}^{(\beta)}(p')$ in (B.7) and $y = \bar{x}^{(\beta)}(p)$ in (B.8) and adding the two equations we obtain

$$\begin{aligned} &\sum_{i\in[R]} \left(c_i(\bar{x}_i^{(\beta)}(p), p) - c_i(\bar{x}_i^{(\beta)}(p'), p')\right)\left(\bar{x}_i^{(\beta)}(p') - \bar{x}_i^{(\beta)}(p)\right) \\ &\quad + \sum_{i\in[R]} \left(\frac{1}{\beta}\ln(\bar{x}_i^{(\beta)}(p)) - \ln(\bar{x}_i^{(\beta)}(p'))\right)\left(\bar{x}_i^{(\beta)}(p') - \bar{x}_i^{(\beta)}(p)\right) \geq 0. \end{aligned}$$

From the fact that prices enter additively, we have

$$\begin{aligned} &\sum_{i\in[R]} \left(\ell_i(\bar{x}_i^{(\beta)}(p)) - \ell_i(\bar{x}_i^{(\beta)}(p'))\right)\left(\bar{x}_i^{(\beta)}(p') - \bar{x}_i^{(\beta)}(p)\right) \\ &\quad + \sum_{i\in[R]} \left(p_i - p_i' + \frac{1}{\beta}\ln(\bar{x}_i^{(\beta)}(p)) - \ln(\bar{x}_i^{(\beta)}(p'))\right)\left(\bar{x}_i^{(\beta)}(p') - \bar{x}_i^{(\beta)}(p)\right) \geq 0. \end{aligned}$$

This gives

$$
\begin{aligned}
&\left\langle \bar{x}^{(\beta)}(p) - \bar{x}^{(\beta)}(p'), p - p' \right\rangle \\
&\leq \left\langle \ell(\bar{x}^{(\beta)}(p)) - \ell(\bar{x}^{(\beta)}(p')), \bar{x}^{(\beta)}(p') - \bar{x}^{(\beta)}(p) \right\rangle \\
&\quad + \frac{1}{\beta} \left\langle \ln(\bar{x}^{(\beta)}(p)) - \ln(\bar{x}^{(\beta)}(p')), \bar{x}^{(\beta)}(p') - \bar{x}^{(\beta)}(p) \right\rangle \\
&< 0,
\end{aligned}
$$

where the last inequality follows due to the monotonicity of the cost function and natural logarithm.

Next, we show that $\bar{x}^{(\beta)}(p)$ is a differentiable function of p. Recall the notation $\mathbf{C}(x) \coloneqq \mathsf{diag}\left(\left(\frac{d\ell_i(x)}{dx}\right)_{i=1}^{R}\right)$ and $\mathbf{M}(x) \coloneqq \mathsf{diag}\left(\left(x_i \frac{d\ell_i(x)}{dx}\right)_{i=1}^{R}\right)$.

To show the differentiability of $\bar{x}^{(\beta)}(p)$ with respect to p, we invoke the implicit function theorem for the map $g(x,p) = x - h(x,p)$. Note that for any fixed p, $\bar{x}^{(\beta)}(p)$ is the zero of $g(\cdot, p)$. To satisfy the hypothesis for the implicit function theorem, for any p we compute the Jacobian of the function $g(x,p)$ with respect to x and evaluate it at $\bar{x}^{(\beta)}(p)$. Using Lemma B.0.1 the Jacobian is given by:

$$
\begin{aligned}
\nabla_x g(\bar{x}^{(\beta)}(p), p) &= I - \nabla_x h(\bar{x}^{(\beta)}(p), p) \\
&= I - \frac{\mu}{\lambda} \beta \bar{x}^{(\beta)}(p) \bar{x}^{(\beta)}(p)^\top \mathbf{C}(\bar{x}^{(\beta)}(p)) + \beta \mathbf{M}\left(\bar{x}^{(\beta)}(p)\right) \\
&= \left(\left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right) - \beta v \mathbb{1}^\top\right) \mathbf{M}(\bar{x}^{(\beta)}(p))
\end{aligned}
$$

where $v \coloneqq \frac{\mu}{\lambda} \bar{x}^{(\beta)}(p)$. Note that from (2.8) we have $\mathbb{1}^\top v = 1$. We claim that $\det(\nabla_x g(\bar{x}^{(\beta)}(p), p)) \neq 0$. Indeed,

$$
\begin{aligned}
\det(\nabla_x g(\bar{x}^{(\beta)}(p), p)) &= \det\left(\mathbf{M}(\bar{x}^{(\beta)}(p)\right) \det\left(\left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right) - \beta v \mathbb{1}^\top\right) \\
&= \det\left(\mathbf{M}(\bar{x}^{(\beta)}(p)\right) \cdot \det\left(I - \beta v \mathbb{1}^\top \left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right)^{-1}\right) \\
&\quad \cdot \det\left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right)^{-1}
\end{aligned}
$$

It is sufficient to show that $\det\left(I - \beta v \mathbb{1}^\top \left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right)^{-1}\right) \neq 0$. Using the Sherman-Morrison formula [17] it is necessary and sufficient to show that $\beta \mathbb{1}^\top \left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right)^{-1} v \neq$

1. We show this by contradiction. Suppose $\beta \mathbb{1}^\top \left(\mathbf{M}^{-1}(\bar{x}^{(\beta)}(p)) + \beta I\right)^{-1} v = 1$. Then

$$\begin{aligned} \implies \quad & \sum_{i=1}^{R} v_i \frac{\beta}{\frac{1}{m_i} + \beta} = 1 \\ \implies \quad & \sum_{i=1}^{R} v_i \frac{\beta}{\frac{1}{m_i} + \beta} = \sum_{i=1}^{R} v_i \\ \implies & \sum_{i=1}^{R} v_i \frac{\frac{1}{m_i}}{\beta + \frac{1}{m_i}} = 0 \end{aligned}$$

This is not possible as the terms on the left hand side are all positive.

The claim that $\frac{\partial \bar{x}_i^{(\beta)}(p)}{\partial p_i} < 0$ also follows from the monotonicity of $\bar{x}^{(\beta)}(p)$. Indeed, if we choose two price profiles that differ only at one index then the monotonicity property of $\bar{x}^{(\beta)}(p)$ ensures $\frac{\partial \bar{x}_i^{(\beta)}(p)}{\partial p_i} < 0$. This completes the proof.

B.0.3 Proof of Lemma 2.2.6

We shall first show the existence and then prove the uniqueness of $\bar{p}$.

For any $p \in \mathbb{R}^R, i \in [R]$, define $z_i(p) = \bar{x}^{(\beta)}(p) \frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx}$. The equilibrium price $\bar{p}$ is then solution to the fixed point equation $p = z(p)$ where $z(p) = [z_i(p)]_{i \in [R]}$. Define a set $K = \{y \in \mathbb{R}^R : y \geq 0, \|y\|_1 \leq \frac{\lambda}{\mu} \max_{i \in [R]} \frac{dc_i\left(\frac{\lambda}{\mu}\right)}{dx}\}$. Observe that $z(\cdot)$ maps the convex compact set K to itself. Therefore, Brouwer's fixed point theorem [41] guarantees the existence of fixed point $\bar{p}$.

Now, we prove the uniqueness of the fixed point $\bar{p}$ satisfying (2.10). We shall prove this via a contradiction argument. Assume that there are two distinct price profiles $\bar{p}, \bar{p}' \in \mathbb{R}^R$ that both satisfy the fixed point equation (2.10), therefore for every $i \in [R]$:

$$\begin{aligned} \bar{p}_i &= \bar{x}_i^{(\beta)}(\bar{p}) \frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} \\ \bar{p}_i' &= \bar{x}_i^{(\beta)}(\bar{p}') \frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}'))}{dx} \end{aligned}$$

Taking the difference we get:

$$
\begin{aligned}
\bar{p}_i - \bar{p}'_i &= \bar{x}_i^{(\beta)}(\bar{p})\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} - \bar{x}_i^{(\beta)}(\bar{p}')\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}'))}{dx}, \\
&= \bar{x}_i^{(\beta)}(\bar{p})\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} - \bar{x}_i^{(\beta)}(\bar{p}')\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} \\
&\qquad + \bar{x}_i^{(\beta)}(\bar{p}')\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} - \bar{x}_i^{(\beta)}(\bar{p}')\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}'))}{dx}, \\
&= \left(\bar{x}_i^{(\beta)}(\bar{p}) - \bar{x}_i^{(\beta)}(\bar{p}')\right)\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} \\
&\qquad + \bar{x}_i^{(\beta)}(\bar{p}')\left(\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} - \frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}'))}{dx}\right),
\end{aligned}
$$

for every $i \in [R]$. Multiplying $(\bar{x}_i^{(\beta)}(\bar{p}) - \bar{x}_i^{(\beta)}(\bar{p}'))$ in the preceding equation we obtain

$$
\begin{aligned}
(\bar{x}_i^{(\beta)}(\bar{p}) - \bar{x}_i^{(\beta)}(\bar{p}'))(\bar{p}_i - \bar{p}'_i) &= \left(\bar{x}_i^{(\beta)}(\bar{p}) - \bar{x}_i^{(\beta)}(\bar{p}')\right)^2 \frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} \\
+ \bar{x}_i^{(\beta)}(\bar{p}')&\left(\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} - \frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}'))}{dx}\right)(\bar{x}_i^{(\beta)}(\bar{p}) - \bar{x}_i^{(\beta)}(\bar{p}')),
\end{aligned}
\tag{B.9}
$$

for every $i \in [R]$. Convexity of the edge cost function ensures that right hand side of (B.9) is always non-negative for every $i \in [R]$. Summing over $i \in [R]$ and using Lemma 2.2.5 we obtain:

$$
0 > \left\langle \bar{x}_i^{(\beta)}(\bar{p}) - \bar{x}_i^{(\beta)}(\bar{p}'), \bar{p} - \bar{p}' \right\rangle \geq 0.
$$

which contradicts our original hypothesis that there are two distinct price profiles satisfying (2.10).

B.0.4 Proof of Lemma 2.2.7

Note that for any $i \in [R]$

$$
\bar{p}_i = \bar{x}_i^{(\beta)}(\bar{p})\frac{d\ell_i(\bar{x}_i^{(\beta)}(\bar{p}))}{dx} \tag{B.10}
$$

Let

$$
x^\star = \arg\min_{y \in F}\left(\sum_{i \in [R]} y_i \ell_i(y_i) + \frac{1}{\beta}\sum_{i \in [R]} y_i \log(y_i)\right),
$$

where $F = \{y : \sum_{i \in [R]} y_i = \frac{\lambda}{\mu}\}$ is the feasible set. We claim that $x^\star = \bar{x}^{(\beta)}(\bar{p})$.

To show this claim, we first note that the above problem is a strictly convex optimization problem. The necessary and sufficient conditions for constrained optimality ensures that for any $y \in F$:

$$\sum_{i\in[R]} \left(\ell_i(x_i^\star) + x_i^\star \frac{d\ell_i(x_i^\star)}{dx} + \frac{1}{\beta}\left(1+\log(x_i^\star)\right)\right)(y_i - x_i^\star) \geq 0. \tag{B.11}$$

Now recall that $\bar{x}^{(\beta)}(\bar{p})$ satisfies

$$\bar{x}^{(\beta)}(\bar{p}) = \arg\min_{y\in F} \sum_{i\in[R]} \int_0^{y_i} \left(\ell_i(z) + \bar{p}_i\right) \mathsf{d}z + \frac{1}{\beta}\sum_{i\in[R]} y_i \log(y_i).$$

Note that we have already noted in proof of Lemma 2.2.4 that the above optimization problem is strictly convex. The constrained optimality conditions ensure that for any $y \in F$:

$$\sum_{i\in[R]} \left(\ell_i(\bar{x}_i^{(\beta)}(\bar{p})) + \bar{p}_i + \frac{1}{\beta}\left(1+\log(\bar{x}_i^{(\beta)}(\bar{p}))\right)\right)\left(y_i - \bar{x}_i^{(\beta)}(\bar{p})\right) \geq 0. \tag{B.12}$$

Note that using (B.10) in (B.12) we obtain that for any $y \in F$:

$$\sum_{i\in[R]} \left(\ell_i(\bar{x}) + \bar{x}_i \frac{d\ell_i(\bar{x}_i)}{dx} + \frac{1}{\beta}\left(1+\log(\bar{x}_i)\right)\right)(y_i - \bar{x}_i) \geq 0, \tag{B.13}$$

where we have used $\bar{x} = \bar{x}^{(\beta)}(\bar{p})$ to simplify the expression. Comparing expression (B.11) and (B.13) we conclude that $x^\star = \bar{x}^{(\beta)}(\bar{p})$. This concludes the proof.

B.0.5 Proof of Lemma 2.2.9

We prove the claims in the lemma sequentially.

Proof of **(C1)**: Using the independence of sequence $\zeta(n), \xi_i(n)$ we see that the conditional expection of $M_i(n+1)$ conditioned on the filtration $\mathcal{F}_n$ is zero. That is,

$$\mathbb{E}[M_i(n+1)|\mathcal{F}_n] = h_i(X_i(n), P_i(n))\left(\mathbb{E}[\zeta(n+1)] - \lambda\right) - X_i(n)\left(\mathbb{E}[\xi_i(n+1) - \mu]\right) = 0.$$

Proof of **(C2)**: Recall that we assume that the incoming loads and outgoing fraction have finite support. That is, for every $i \in [R]$ and n we have $\zeta(n) \in [\underline{\lambda}, \bar{\lambda}] \subset (0, \infty)$ and $\xi_i(n) \in [\underline{\mu}, \bar{\mu}] \subset (0, 1)$. Note that $X_i(0) \geq 0$ for all $i \in [R]$ as it is impractical to have negative load on the network; therefore our update reflects the physical constraints of the network. Under this assumption, the discrete-time stochastic update (2.4) can be written as

$$
\begin{aligned}
X_i(n+1) &= (1-\xi(n+1))X_i(n) + \frac{\exp(-\beta c_i(X_i(n), P_i(n)))}{\sum_{j\in[R]} \exp(-\beta c_j(X_j(n), P_j(n)))}\zeta(n), \\
&\leq (1-\underline{\mu})X_i(n) + \zeta(n), \\
&\leq (1-\underline{\mu})X_i(n) + \bar{\lambda}, \\
&\leq (1-\underline{\mu})^n X_i(0) + \frac{\bar{\lambda}}{\underline{\mu}}.
\end{aligned}
$$

The preceding inequalities establish that $\mathbb{E}[\|X_i(n)\|^2] < +\infty$. Consequently, it also ensure that $\mathbb{E}\left[X_i(n)\frac{d\ell_i(n)}{dx}\right] < +\infty$ and is independent of n, which in turn, from (Update-P) leads to $\mathbb{E}[\|P_i(n)\|^2] < +\infty$.

Proof of **(C3)**: We see that:

$$
\begin{aligned}
\mathbb{E}[|M_i(n+1)|^2|\mathcal{F}_n] &= \mathbb{E}[|h_i(X_i(n), P_i(n))(\zeta(n+1)-\lambda) - X_i(n)\left(\xi_i(n+1)-\mu\right)|^2|\mathcal{F}_n] \\
&\leq \mathbb{E}[2|h_i(X_i(n), P_i(n))(\zeta(n+1)-\lambda)|^2 + 2|X_i(n)\left(\xi_i(n+1)-\mu\right)|^2|\mathcal{F}_n] \\
&\leq \mathbb{E}[2\frac{\lambda}{\mu}|(\zeta(n+1)-\lambda)|^2 + 2|X_i(n)\left(\xi_i(n+1)-\mu\right)|^2|\mathcal{F}_n] \\
&\leq K_1 + K_2\|X_i(n)\|^2 \\
&\leq K\left(1+|X_i(n)|^2\right) < +\infty
\end{aligned}
$$

where $K_1 = \frac{2\lambda}{\mu}|\bar{\lambda}-\underline{\lambda}|^2$, $K_2 = 2|\bar{\mu}-\underline{\mu}|^2$ and $K = \max(K_1, K_2)$. To obtain the preceding bound we used $h_i(X_i(n), P_i(n)) \leq \frac{\lambda}{\mu}$.

Proof of **(C4)**: To prove this result we shall use Theorem A.0.2. For that purpose we need to check conditions Theorem A.0.2(P-i)-(P-iv). The condition (P-i) is satisfied if we show that $\frac{\partial h_i(x,p)}{\partial x_j} \geq 0$. Indeed, by definition

$$
\frac{\partial h_i(x,p)}{\partial x_j} = \frac{\lambda}{\mu}\exp(-\beta c_i(x_i,p_i))\frac{\beta\exp(-\beta c_j(x_j,p_j))\frac{\partial c_j(x_j,p_j)}{\partial x_j}}{\left(\sum_{j\in[R]}\exp\left(-\beta c_j(x_j,p_j)\right)\right)^2} > 0.
$$

Furthermore, note that condition (P-ii) is also satisfied because the Jacobian matrix has all the off-diagonal terms positive and is therefore irreducible. Moreover, it follows from (2.5) that for any $x, p \in \mathbb{R}^R$, $h_i(x,p) \in \left(0, \frac{\lambda}{\mu}\right)$ and therefore (P-iii)-(P-iv) are also satisfied. This completes the proof.

Proof of **(C5)**: We shall now show that (2.16) satisfies conditions (P-i)-(P-iv) in Theorem A.0.2. This ensures global convergence to the equilibrium set which is guaranteed to be singleton by Lemma 2.2.6. For better presentation, for any $i, j \in [R]$, we denote the derivative of $\bar{x}_i^{(\beta)}(p)$ with respect to p_j by $\nabla_j \bar{x}^{(\beta)}(p)$[1]. Using (2.8) for any $i \neq j$ we have

$$\exp(-\beta c_j(\bar{x}_j^{(\beta)}(p), p_j))\bar{x}_i^{(\beta)}(p) = \exp(-\beta c_i(\bar{x}_i^{(\beta)}(p), p_i))\bar{x}_j^{(\beta)}(p) \tag{B.14}$$

Taking derivative of the above equation with respect to p_k for $k \neq i \neq j$:

$$\begin{aligned}
&\exp(-\beta c_j(\bar{x}_j^{(\beta)}(p), p_j))\nabla_k \bar{x}_i^{(\beta)}(p) \\
&\quad + \bar{x}_i^{(\beta)}(p)\exp(-\beta c_j(\bar{x}_j^{(\beta)}(p), p_j))(-\beta \nabla_x c_j(\bar{x}_j^{(\beta)}(p), p_j)\nabla_k \bar{x}_j^{(\beta)}(p)) \\
&= \exp(-\beta c_i(\bar{x}_i^{(\beta)}(p), p_i))\nabla_k \bar{x}_j^{(\beta)}(p) \\
&\quad + \bar{x}_j^{(\beta)}(p)\exp(-\beta c_i(\bar{x}_i^{(\beta)}(p), p_i))(-\beta \nabla_x c_i(\bar{x}_i^{(\beta)}(p), p_i)\nabla_k \bar{x}_i^{(\beta)}(p)).
\end{aligned}$$

Collecting similar terms together we get

$$\begin{aligned}
&\nabla_k \bar{x}_i^{(\beta)}(p)\left(\exp(-\beta c_j(\bar{x}_j^{(\beta)}(p), p_j))\right) \\
&\quad + \nabla_k \bar{x}_i^{(\beta)}(p)\left(\bar{x}_j^{(\beta)}(p)\exp(-\beta c_i(\bar{x}_i^{(\beta)}(p), p_i))(\beta \nabla_x c_i(\bar{x}_i^{(\beta)}(p), p_i))\right) \\
&= \nabla_k \bar{x}_j^{(\beta)}(p)\left(\exp(-\beta c_i(\bar{x}_i^{(\beta)}(p), p_i))\right) \\
&\quad + \nabla_k \bar{x}_j^{(\beta)}(p)\left(\bar{x}_i^{(\beta)}(p)\exp(-\beta c_j(\bar{x}_j^{(\beta)}(p), p_j))(\beta \nabla_x c_j(\bar{x}_j^{(\beta)}(p), p_j))\right).
\end{aligned}$$

This implies for $i \neq j \neq k$ and for any $p \in \mathbb{R}^R$ we have

$$\nabla_k \bar{x}_i^{(\beta)}(p) \cdot \nabla_k \bar{x}_j^{(\beta)}(p) > 0. \tag{B.15}$$

Moreover, by definition of fixed point in (2.8) we have the constraint that

$$\sum_{l \in [R]} \bar{x}_l^{(\beta)}(p) = \frac{\lambda}{\mu}.$$

Taking the derivative with respect to p_k of the above equation we obtain

$$\sum_{l \neq k} \nabla_k \bar{x}_l^{(\beta)}(p) = -\nabla_k \bar{x}_k^{(\beta)}(p) > 0, \tag{B.16}$$

where the last inequality follows from Lemma 2.2.5. Equation (B.16) in conjunction with (B.15) implies that $\nabla_k \bar{x}_i^{(\beta)}(p) > 0$ for all $i \neq k$. This ensures satisfaction of (P-i)-(P-ii). The requirement (P-iii)-(P-iv) is also satisfied as for any $p \in \mathbb{R}^R$, $\bar{x}^{(\beta)}(p) \in \left(0, \frac{\lambda}{\mu}\right)$. This completes the proof.

[1]Note that in Lemma 2.2.5, we established that $\bar{x}^{(\beta)}(p)$ is continuously differentiable

B.0.6 Proof of Lemma 2.2.10

Note that to invoke the results from two timescale stochastic approximation theory [5], in addition to Lemma 2.2.9 we also need to ensure the following

(i) the function $\bar{x}^{(\beta)}(p)$ is Lipschitz;

(ii) the function $g(x,p) \coloneqq h(x,p) - x$, which is the vector field in (2.15), is Lipschitz;

(iii) the function $r_i(p) \coloneqq \bar{x}_i^{(\beta)}(p)\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} - p_i$, which is the vector field in (2.16), is Lipschitz for all $i \in [R]$.

If the above conditions hold, then [5, Chapter 9] ensures that Theorem 2.2.3 hold.

Note that (i) holds due the fact that $\bar{x}^{(\beta)}(p) \in (0, \lambda/\mu)$ and Lemma 2.2.5 where we established that it is continuously differentiable. Moreover, (ii) holds due to Lemma B.0.1. At last, to show (iii) we note that for any $p, p' \in \mathbb{R}^R$ and $i \in [R]$:

$$
\begin{aligned}
&\|r_i(p) - r_i(p')\| = \Big\|\bar{x}_i^{(\beta)}(p)\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} - \bar{x}_i^{(\beta)}(p')\frac{d\ell_i(\bar{x}_i^{(\beta)}(p'))}{dx}\Big\| \\
&\leq \Big\|\bar{x}_i^{(\beta)}(p)\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} - \bar{x}_i^{(\beta)}(p')\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx}\Big\| \\
&\quad + \Big\|\bar{x}_i^{(\beta)}(p')\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} - \bar{x}_i^{(\beta)}(p')\frac{d\ell_i(\bar{x}_i^{(\beta)}(p'))}{dx}\Big\| \\
&\leq \Big|\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx}\Big|\|\bar{x}_i^{(\beta)}(p) - \bar{x}_i^{(\beta)}(p')\| + |\bar{x}_i^{(\beta)}(p')|\Big\|\frac{d\ell_i(\bar{x}_i^{(\beta)}(p))}{dx} - \frac{d\ell_i(\bar{x}_i^{(\beta)}(p'))}{dx}\Big\| \\
&\leq K_1\bar{L}\|p - p'\| + K_2\tilde{L}\|p - p'\|,
\end{aligned}
$$

where $K_1 = \max_{i\in[R],x\in[0,\lambda/\mu]}\big|\frac{d\ell_i(x)}{dx}\big|$, $\bar{L}$ is the Lipschitz constant for $\bar{x}^{(\beta)}(\cdot)$, $K_2 = \lambda/\mu\bar{L}$ and $\tilde{L}$ is the Lipschitz constant for $\frac{d\ell_i(x)}{dx}$ when $x \in [0, \lambda/\mu]$.

B.0.7 Proof of Proposition 3.2.1

Proof. We provide a detailed proof for the setting of atomic games as the proof for the non-atomic game follows similarly.

Atomic game G: We show that $P^\dagger$ is non-empty. That is, there exists $p^\dagger$ such that $e(x^*(p^\dagger)) = p^\dagger$. Define a function $\theta(p) = e(x^*(p))$. Thus, the remaining proof is based on application of Brouwer's fixed point theorem to show existence of the fixed points of function $\theta(\cdot)$.

We note that $\theta(p)$ is a continuous function based on the setup presented in Sec 2.1. Furthermore, let's define $K \coloneqq \{\theta(p) : p \in \mathbb{R}^{|\mathcal{I}|}\} \subset \mathbb{R}^{|\mathcal{I}|}$. We claim that the set K is compact. Indeed, this follows by two observations. First, the externality function $e(\cdot)$ is continuous. Second, the range of the function $x^*(\cdot)$ is X which is a compact space. These two observations ensure that $\theta(p) = e(x^*(p))$ is a bounded function. Let $\tilde{K} \coloneqq \mathsf{conv}(K)$ be the convex hull of K, which in turn is also a compact set. Let's denote the restriction of function θ on the set $\tilde{K}$ as $\theta_{|\tilde{K}} : \tilde{K} \longrightarrow \tilde{K}$ where $\theta_{|\tilde{K}}(p) = \theta(p)$ for all $p \in \tilde{K}$. We note that $\theta_{|\tilde{K}}$ a is continuous function from a convex compact set to itself and therefore Brouwer's fixed point theorem ensures that there exists $p^\dagger \in \tilde{K}$ such that $p^\dagger = \theta_{|\tilde{K}}(p^\dagger) = \theta(p^\dagger)$. This concludes the proof about existence of $p^\dagger$.

Next, we show that incentive $p^\dagger$ aligns Nash equilibrium with social optimality (i.e. for any $p^\dagger \in P^\dagger$, $x^*(p^\dagger) = x^\dagger$). Fix $p^\dagger \in P^\dagger$. For every $i \in \mathcal{I}$ we have $p_i^\dagger = e_i(x^*(p^\dagger))$. This implies $D_{x_i}\ell_i(x^*(p^\dagger)) + p_i^\dagger = D_{x_i}\Phi(x^*(p^\dagger))$ for every $i \in \mathcal{I}$. This implies

$$J(x^*(p^\dagger), p^\dagger) = \nabla\Phi(x^*(p^\dagger)). \tag{B.17}$$

Next, from (3.2) we know that $x^*(p^\dagger)$ is a Nash equilibrium if and only if

$$\langle J(x^*(p^\dagger), p^\dagger), x - x^*(p^\dagger)\rangle \geq 0, \quad \forall\, x \in X. \tag{B.18}$$

Using (B.17) and (B.18) the following holds:

$$\langle \nabla\Phi(x^*(p^\dagger)), x - x^*(p^\dagger)\rangle \geq 0, \quad \forall\, x \in X. \tag{B.19}$$

Comparing (B.19) with (3.4) we note that $x^*(p^\dagger)$ is the minimizer of social cost function Φ. This implies $x^*(p^\dagger) = x^\dagger$ as $x^\dagger$ is the unique minimizer of social cost function Φ.

□

B.0.8 Proof of Proposition 3.2.2

Proof. The proof is based on a contradiction argument.

(i) We make the following observation which are central to the proof:

(O1) We note that if $x^*(p) \in \mathsf{int}(X)$ for every p then the variational inequality characterization (3.2) implies that $J(x^*(p), p) = 0$ for every p. As a result the externality function (3.5) becomes $e(x^*(p)) = \nabla\Phi(x^*(p)) + p$.

(O2) The strict convexity of the social cost function implies that

$$\langle \nabla\Phi(x) - \nabla\Phi(y), x - y\rangle > 0, \quad \forall\, x, y \in X \text{ such that } x \neq y$$

Suppose there exists two distinct elements $p^\dagger, q^\dagger \in P^\dagger$. We claim that $x^*(p^\dagger) \neq x^*(q^\dagger)$. Indeed, using **(O1)** and (3.3) the following holds:

$$\begin{aligned} p_i^\dagger &= -D_{x_i}\ell(x^*(p^\dagger)), \quad \forall\, i \in \mathcal{I} \\ q_i^\dagger &= -D_{x_i}\ell(x^*(q^\dagger)), \quad \forall\, i \in \mathcal{I}. \end{aligned} \tag{B.20}$$

If $x^*(p^\dagger) = x^*(q^\dagger)$ then (B.20) implies $p^\dagger = q^\dagger$, but these are assumed to be distinct. Thus in the following proof we assume $x^*(p^\dagger) \neq x^*(q^\dagger)$.

We note from **(O1)** that

$$0 = \nabla\Phi(x^*(p^\dagger)), \quad 0 = \nabla\Phi(x^*(q^\dagger)). \tag{B.21}$$

Substracting the two expressions in (B.21) and taking inner product with $x^*(p^\dagger) - x^*(q^\dagger)$ we see that

$$0 = \left\langle x^*(p^\dagger) - x^*(q^\dagger), \nabla\Phi(x^*(p^\dagger)) - \nabla\Phi(x^*(q^\dagger))\right\rangle \tag{B.22}$$

We arrive at a contradiction by noting that $x^*(p^\dagger) \neq x^*(q^\dagger)$ and **(O2)** imply that RHS is strictly positive.

(ii) To begin the proof we define $D\ell(x) = (D_{x_i}\ell_i(x))_{i\in\mathcal{I}}$. Under this notation, we have $e(x) = \nabla\Phi(x) - D\ell(x)$. The proof is based on the following observations:

(O3) We claim that $\langle x^*(p_1) - x^*(p_2), p_1 - p_2\rangle < 0$ for any two distinct incentives $p_1 \neq p_2$. Indeed, from the variational inequality characterization of Nash equilibrium (3.2) we know that

$$\begin{aligned} \langle D\ell(x^*(p_1)) + p_1, x_1 - \tilde{x}^*(p_1)\rangle \geq 0, \quad \forall x_1 \in X \\ \langle D\ell(x^*(p_2)) + p_2, x_2 - \tilde{x}^*(p_2)\rangle \geq 0, \quad \forall x_2 \in X \end{aligned}$$

Picking $x_1 = \tilde{x}^*(p_2)$ and $x_2 = \tilde{x}^*(p_1)$, and adding the two inequalities in preceding equation we obtain

$$\langle x^*(p_1) - x^*(p_2), p_1 - p_2 \rangle \leq -\langle D\ell(x^*(p_1)) - D\ell(x^*(p_2)), x^*(p_1) - x^*(p_2) \rangle \leq 0$$

where the last inequality follows due to the convexity of ℓ.

We prove the uniqueness by contradiction. Suppose there exists two incentives $p^\dagger, q^\dagger \in P^\dagger$ such that $\tilde{e}(\tilde{x}^*(p^\dagger)) = p^\dagger$ and $\tilde{e}(\tilde{x}^*(q^\dagger)) = q^\dagger$. Then we have

$$\begin{aligned} p^\dagger &= \nabla\Phi(x^*(p^\dagger)) - D\ell(x^*(p^\dagger)) \\ q^\dagger &= \nabla\Phi(x^*(q^\dagger)) - D\ell(x^*(q^\dagger)). \end{aligned}$$

Subtracting the two expressions and taking inner product with $x^*(p^\dagger) - x^*(q^\dagger)$ we have

$$\left\langle x^*(p^\dagger) - x^*(q^\dagger), p^\dagger - q^\dagger \right\rangle = \left\langle x^*(p^\dagger) - x^*(q^\dagger), e(x^*(p^\dagger)) - e(x^*(q^\dagger)) \right\rangle > 0.$$

But from **(O3)** we see that we arrive at a contradiction as $\left\langle x^*(p^\dagger) - x^*(q^\dagger), p^\dagger - q^\dagger \right\rangle \leq 0$.

□

B.0.9 Proof of Theorem 3.2.3

Proof. To ensure the convergence of (x_k, p_k) to the fixed point $(x^\dagger, p^\dagger)$ of (x-update)-(p-update), we exploit the timescale separation introduced due to Assumption 3.1.2. The proof is based on two-timescale dynamical systems theory described in [4]. Due to this timescale separation the strategy update evolves faster than the incentive update. This allows us to appropriately decouple the strategy and incentive update and analyze them separately.

Note that we can equivalently write (x-update)-(p-update) as follows:

$$\begin{aligned} x_{k+1} &= x_k + \gamma_k \left(f(x_k, p_k) - x_k \right) \\ p_{k+1} &= p_k + \gamma_k \left(\frac{\beta_k}{\gamma_k} \left(e(x_k) - p_k \right) \right), \end{aligned} \tag{B.23}$$

where $\lim_{k \longrightarrow \infty} \frac{\beta_k}{\gamma_k} = 0$ and $\lim_{k \longrightarrow \infty} \gamma_k = 0$. From two timescale dynamical systems theory we know that under Assumption 3.1.3 the tuple (x_k, p_k) converges to the set $\{(x^*(p), p) : p \in \mathbb{R}^{|\mathcal{I}|}\}$. Thus for sufficiently large values of k, the update x_k closely tracks $x^*(p_k)$. Therefore, we consider the following update to analyze the convergence of the slow incentive update (p-update):

$$\mathbf{p}_{k+1} = (1-\beta_k)\mathbf{p}_k + \beta_k e(x^*(\mathbf{p}_k)). \tag{B.24}$$

Since the step sizes $\{\beta_k\}$ are asymptotically going to zero and is non-summable (Assumption 3.1.2-(i)) we can approximate the updates in (B.24) by the following continuous-time dynamical system :

$$\dot{\mathbf{p}}(t) = e(x^*(\mathbf{p}(t))) - \mathbf{p}(t), \tag{B.25}$$

Convergence of discrete-time updates (x-update)-(p-update) then hold if the flow of (B.25) globally converges to $P^\dagger$.

Requirements (C1) in Theorem 3.2.3 is a sufficient condition for convergence of the trajectories of (B.25) to the set $P^\dagger$. This condition is based on cooperative dynamical systems theory [19]. On the other hand requirement (C2) in Theorem 3.2.3 ensures convergence the trajectories of (B.25) to the set $P^\dagger$ by demanding existence of a Lyapunov function [41] that is strictly positive everywhere except at $P^\dagger$ and decreases along the flow of (B.25).

□

B.0.10 Proof of Proposition 3.3.1

Proof. We first show that the set $P^\dagger$ is singleton for the setup in Sec 3.3.1[2]. Then we show that the dynamic update (x_k, p_k) corresponding to (3.16a)-(3.16b) converges to the social optimality $(x^\dagger, p^\dagger)$.

Note that any element $p^\dagger \in P^\dagger$ should satisfy $p^\dagger = e(x^*(p^\dagger)) = \xi + Awx^*(p^\dagger)$. Moreover, from (3.15) we know that $x^*(p^\dagger) = Awx^*(p^\dagger) - p^\dagger$. Succintly writing the preceding two relations in matrix form gives us:

$$\underbrace{\begin{bmatrix} I & -Aw \\ I & I - Aw \end{bmatrix}}_{\Gamma} \begin{bmatrix} p^\dagger \\ x^*(p^\dagger) \end{bmatrix} = \begin{bmatrix} \xi \\ 0 \end{bmatrix}.$$

[2] Note that we cannot directly use Proposition 3.2.1 as that requires compactness of strategy space.

We claim that Γ is an invertible matrix[3] with the inverse as follows:

$$\Gamma^{-1} = \begin{bmatrix} I - Aw & Aw \\ -I & I \end{bmatrix}$$

Thus $(x^*(p^\dagger), p^\dagger)$ exists and is unique. Moreover $p^\dagger = (I - Aw)\xi$ and $x^*(p^\dagger) = -\xi = x^\dagger$.

Next, to ensure that the dynamic update (x_k, p_k) corresponding to (3.15)-(3.16b) converges to the fixed point $(x^\dagger, p^\dagger)$ we use Theorem 3.2.3. It is sufficient to show that Assumption 3.1.3 and condition **(C2)** hold in order to use the results from Theorem 3.2.3 directly.

First, we show that Assumption 3.1.3 hold. That is, for any fixed incentive update $(p_k) \equiv p$ the strategy update satisfies $\lim_{k \longrightarrow \infty} x_k = x^*(p)$. Indeed, due to Assumption 3.1.2, the convergence properties of discrete time updates can be obtained by analysing the corresponding continuous time dynamical system. That is, we consider the following continuous time dynamical system corresponding to the strategy update:

$$\dot{\mathbf{x}}(t) = -(I - Aw)\mathbf{x}(t) - p. \tag{B.26}$$

We note that the trajectories of (B.26) satisfy $\lim_{t \longrightarrow \infty} \mathbf{x}(t) = x^*(p)$. This is due to the assumption that $-(I - Aw)$ is Hurwitz[4] [8]. Thus Assumption 3.1.3 holds.

Next, we show that condition **(C2)** is satisfied which then fulfils all the requirement of Theorem 3.2.3. We claim that the function $V(p) = (p - p^\dagger)^\top L(p - p^\dagger)$ satisfies **(C2)** where L^5 is a symmetric positive definite matrix that satisfies the following condition:

$$(I - Aw)^{-\top} M + M(I - Aw)^{-1} = I. \tag{B.27}$$

[3]Invertibility of $I - Aw$ is a necessary condition for invertiblility of Γ.
[4]A matrix A is called Hurwitz if $\mathsf{spec}(A) \subset \mathbb{C}^\circ_-$.
[5]Note that the existence of such a matrix L is guaranteed from Lyapunov's theorem [8] as $\mathsf{spec}(I - Aw) \subset \mathbb{C}^\circ_+$.

Indeed, $V(p^\dagger) = 0$ and since L is a positive definite matrix, this means $V(p) > 0$ for all $p \neq p^\dagger$. Furthermore, we compute

$$\begin{aligned}
\nabla V(p)^\top (e(x^*(p)) - p) &= 2(p-p^\dagger)^\top L(e(x^*(p)) - p), \\
&\underset{(a)}{=} 2(p-p^\dagger)^\top L\left(\xi + Awx^*(p) - p\right), \\
&= 2(p-p^\dagger)^\top L\left(-x^*(p^\dagger) + Awx^*(p) - p\right), \\
&\underset{(b)}{=} 2(p-p^\dagger)^\top L\left(-x^*(p^\dagger) + x^*(p)\right), \\
&\underset{(c)}{=} -2(p-p^\dagger)^\top L(I-Aw)^{-1}(p-p^\dagger), \\
&\underset{(d)}{=} -(p-p^\dagger)^\top \left(L(I-Aw)^{-1} + (I-Aw)^{-\top} L\right)(p-p^\dagger), \\
&= -(p-p^\dagger)^\top (p-p^\dagger) < 0,
\end{aligned}$$

where (a) is by the definition of externality function (3.5), $(b), (c)$ is by the Nash equilibrium (3.15) and (d) is by (B.27). This completes the proof.

□

B.0.11 Proof of Proposition 3.3.2

Before stating the proof of Proposition 3.3.2 we present the following two results which are crucial in the proof of Proposition 3.3.2. First, we prove the Nash equilibrium takes the form stated in (3.18). Next, we present a technical lemma.

Below, we state the Nash equilibrium in Cournot competition in terms of incentives.

Lemma B.0.2 (Nash equilibrium)**.** *For any given incentive p, the Nash equilibrium is given by*

$$x^*(p) = \frac{\theta - \nu}{\delta(|\mathcal{I}|+1)}\mathbb{1} - \frac{1}{\delta}p + \frac{1}{\delta(|\mathcal{I}|+1)}\mathbb{1}\mathbb{1}^\top p$$

Proof. In the setup of Sec 3.3.2 the variational inequality characterization of Nash equilibrium (3.2) implies that for any given p, $x^*(p)$ is a Nash equilibrium if and only if $J(x^*(p), p) = 0$. Consequently, $x^*(p)$ satisfies the following

$$2x_i^*(p) + \sum_{j \neq i} x_j^*(p) = \frac{\theta - \nu - p_i}{\delta}.$$

Recasting this in the matrix form gives the following:

$$\underbrace{\begin{bmatrix} 2 & 1 & \dots & 1 \\ 1 & 2 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 2 \end{bmatrix}}_{A} \begin{bmatrix} x_1^*(p) \\ x_2^*(p) \\ \dots \\ x_n^*(p) \end{bmatrix} = \begin{bmatrix} \frac{\theta-\nu-p_1}{\delta} \\ \frac{\theta-\nu-p_2}{\delta} \\ \dots \\ \frac{\theta-\nu-p_n}{\delta} \end{bmatrix}. \tag{B.28}$$

Note that $A = I + \mathbb{1}\mathbb{1}^\top$. Furthermore, by the Sherman-Morrison formula:

$$A^{-1} = \frac{1}{|\mathcal{I}|+1} \begin{bmatrix} |\mathcal{I}| & -1 & \dots & -1 \\ -1 & |\mathcal{I}| & \dots & -1 \\ \dots & \dots & \dots & \dots \\ -1 & -1 & \dots & |\mathcal{I}| \end{bmatrix} = I - \frac{1}{|\mathcal{I}|+1}\mathbb{1}\mathbb{1}^\top. \tag{B.29}$$

Therefore, by (B.28) we have :

$$\begin{bmatrix} x_1^*(p) \\ x_2^*(p) \\ \dots \\ x_n^*(p) \end{bmatrix} = \frac{1}{\delta(|\mathcal{I}|+1)} \begin{bmatrix} \theta - \nu - |\mathcal{I}|p_1 + \sum_{j\neq 1} p_j \\ \theta - \nu - |\mathcal{I}|p_2 + \sum_{j\neq 2} p_j \\ \dots \\ \theta - \nu - |\mathcal{I}|p_n + \sum_{j\neq n} p_j \end{bmatrix}.$$

This completes the proof. $\square$

Next, we present a technical lemma that is crucial in the proof of Proposition 3.3.2.

Lemma B.0.3. *Let* $\Gamma = (2\lambda - \delta)I + \delta\mathbb{1}\mathbb{1}^\top$ *and* $\Omega = -\frac{1}{\delta}I + \frac{1}{\delta(|\mathcal{I}|+1)}\mathbb{1}\mathbb{1}^\top$. *If* $\lambda > \delta$ *then* *spec*$(\Gamma\Omega) \subset \mathbb{C}_-^\circ$.

Proof. Note that

$$\begin{aligned} -\Gamma\Omega &= \frac{1}{\delta(|\mathcal{I}|+1)}((2\lambda-\delta)I + \delta\mathbb{1}\mathbb{1}^\top)((|\mathcal{I}|+1)I - \mathbb{1}\mathbb{1}^\top) \\ &= \frac{1}{\delta(|\mathcal{I}|+1)}\left((|\mathcal{I}|+1)(2\lambda-\delta)I - (2\lambda-\delta)\mathbb{1}\mathbb{1}^\top + (|\mathcal{I}|+1)\delta\mathbb{1}\mathbb{1}^\top - |\mathcal{I}|\delta\mathbb{1}\mathbb{1}^\top\right) \\ &= \frac{1}{\delta(|\mathcal{I}|+1)}\left((|\mathcal{I}|+1)(2\lambda-\delta)I - (2\lambda-2\delta)\mathbb{1}\mathbb{1}^\top\right) \end{aligned}$$

From Gershgorin's circle theorem[6] we know that for the result to hold it is sufficient to ensure

$$|(|\mathcal{I}|+1)(2\lambda-\delta) - (2\lambda-2\delta)| > (|\mathcal{I}|-1)|2\lambda-2\delta| \tag{B.31}$$

[6]Gershgorin's circle theorem [45], which says that for a square matrix $A \in \mathbb{R}^{n\times n}$, each eigenvalue of A is contained in at

In fact if $\lambda > \delta$ then (B.31) holds.

$\square$

Finally, we present the proof of Proposition 3.3.2 below:

Proof of Proposition 3.3.2. We show that the set $P^\dagger$ is singleton for the setup in Section 3.3.1[7]. Then we show that the dynamic update (x_k, p_k) corresponding to (3.16a)-(3.16b) converges to social optimality $(x^\dagger, p^\dagger)$.

Note that any element $p^\dagger \in P^\dagger$ should satisfy $p^\dagger = e(x^*(p^\dagger)) = \left((2\lambda - \delta)I + \delta \mathbb{1}\mathbb{1}^\top\right) x^*(p^\dagger)$. Moreover from Lemma B.0.2 we know that $x^*(p^\dagger) = \frac{\theta-\nu}{\delta(|\mathcal{I}|+1)}\mathbb{1} - \frac{1}{\delta}\left(I - \frac{\mathbb{1}\mathbb{1}^\top}{|\mathcal{I}|+1}\right)p^\dagger$. Succinctly writing these two requirements in matrix form gives us:

$$\underbrace{\begin{bmatrix} (2\lambda-\delta)I + \delta\mathbb{1}\mathbb{1}^\top & -I \\ \delta I & I - \frac{1}{|\mathcal{I}|+1}\mathbb{1}\mathbb{1}^\top \end{bmatrix}}_{B} \begin{bmatrix} x^*(p^\dagger) \\ p^\dagger \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{\theta-\nu}{|\mathcal{I}|+1}\mathbb{1} \end{bmatrix}$$

We claim that B is an invertible matrix. Indeed, lower diagonal is an invertible block by (B.29) and the Schur complement of B with respect to that block is $2\lambda I + 2\delta\mathbb{1}\mathbb{1}^\top$ which is also invertible. Thus $(x^*(p^\dagger), p^\dagger)$ exists and is unique.

Next, we use Theorem 3.2.3 to ensure that the dynamic update (x_k, p_k) corresponding to (3.19a)-(3.19b) converges to the fixed point $(x^*(p^\dagger), p^\dagger)$. It is sufficient to show that Assumption 3.1.3 and condition **(C2)** hold. Before checking these conditions we define $\Gamma = (2\lambda - \delta)I + \delta\mathbb{1}\mathbb{1}^\top$ and $\Omega = -\frac{1}{\delta}I + \frac{1}{\delta(|\mathcal{I}|+1)}\mathbb{1}\mathbb{1}^\top$.

First, we show that Assumption 3.1.3 holds. That is, for any fixed incentive $(p_k) = p$, the strategy update satisfies $\lim_{k\to\infty} x_k = x^*(p)$. Indeed, due to Assumption 3.1.2, the convergence properties of discrete time updates can be obtained by analysing the corresponding

least one of the disks:

$$D_i = \{z \in \mathbb{C} : |z - A_{ii}| \le \sum_{j\neq i} |A_{ij}|\} \tag{B.30}$$

where A_{ii} are the diagonal entries of A, and A_{ij} are the off-diagonal entries. In our case, to ensure that $\Gamma\Omega$ has eigenvalues on the open right half plane, we need to ensure that $|A_{ii}| - \sum_{j\neq i}|A_{ij}| > 0$

[7]Note that we cannot directly use Proposition 3.2.1 as that requires compactness of strategy space.

continuous time dynamical system stated below:

$$\dot{\mathbf{x}}_i(t) = -\mathbf{x}_i(t) + \left(\frac{\theta - \delta \sum_{j\neq i} \mathbf{x}_j(t) - \nu - p_i}{2\delta}\right) \tag{B.32}$$

We claim that the trajectories of (B.32) satisfy $\lim_{t\to 0} \mathbf{x}(t) = x^*(p)$. Indeed, the atomic Cournot competition is a potential game with the following potential function for any p:

$$T(\mathbf{x}, p) = -\int_0^{\sum_{i=1}^{|\mathcal{I}|} \mathbf{x}_i} (\theta - \delta z)dz + \frac{\delta}{2}\sum_{i\in\mathcal{I}} x_i^2 + \sum_{i=1}^{|\mathcal{I}|}(\nu + p_i)x_i, \tag{B.33}$$

and (B.32) is the corresponding continuous time best response dynamics. Thus [44, Theorem 2] ensures $\lim_{t\to\infty} \mathbf{x}(t) = x^*(p)$.

Next, we show that condition **(C2)** is satisfied, which fulfills the requirements of Theorem (3.2.3). We claim that the function $V(p) = (p - p^\dagger)L(p - p^\dagger)$ satisfies **(C2)**, where L is a symmetric positive definite matrix that satisfies:

$$(\Gamma\Omega)^\top L + L^\top(\Gamma\Omega) = -I \tag{B.34}$$

Note that the existence of L follows from the Lyapunov theorem [8] as from Lemma B.0.3 we know that $-\Gamma\Omega$ is a Hurwitz matrix.

Indeed, $V(p^\dagger) = 0$ and since L is positive definite, this means $V(p) > 0$ for all $p \neq p^\dagger$. Furthermore, we compute:

$$\begin{aligned}
\nabla V(p)^\top(e(x^*(p)) - p) &= 2(p - p^\dagger)^\top L(e(x^*(p)) - p)\\
&\underset{(a)}{=} 2(p - p^\dagger)^\top L\left(\left((2\lambda - \delta)I + \delta\mathbb{1}\mathbb{1}^\top\right)x^*(p) - p\right)\\
&\underset{(b)}{=} 2(p - p^\dagger)^\top L\left(\left((2\lambda - \delta)I + \delta\mathbb{1}\mathbb{1}^\top\right)(x^*(p) - x^*(p^\dagger)) + p^\dagger - p\right)\\
&= 2(p - p^\dagger)^\top L\Gamma(x^*(p) - x^*(p^\dagger)) - 2(p - p^\dagger)^\top L(p - p^\dagger)\\
&\underset{(c)}{=} 2(p - p^\dagger)^\top L\Gamma\Omega(p - p^\dagger) - 2(p - p^\dagger)^\top L(p - p^\dagger)\\
&\underset{(d)}{=} (p - p^\dagger)^\top(L(\Gamma\Omega) + (\Gamma\Omega)^\top L)(p - p^\dagger) - 2(p - p^\dagger)^\top L(p - p^\dagger)\\
&= -(p - p^\dagger)(2L + I)(p - p^\dagger)
\end{aligned}$$

where (a) is by the definition of the externality function $e(x^*(p))$, (b) is by adding and subtracting $p^\dagger$, (c) is by the definition of the Nash equilibrium $x^*(p)$, and (d) is by (B.34). This completes the proof. □