Game Theoretic Models of Networks Security

by

Stamatios Katsikas

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy in Interdisciplinary

Mathematics and Complexity Science

Centre for Complexity Science

September 2017

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

It has been a unique privilege to work with Vassili Kolokoltsov and Steve Alpern, both for their scientific excellence and their enquiring minds. Without their eager guidance, generous assistance and understanding, this thesis would not have been completed. I will always be grateful to both of them for that. I would like to thank the Centre for Complexity Science directorate for the opportunity to join the centre as a PhD student, and the administrative staff for the support throughout my term. I would also like to acknowledge support of the Engineering and Physical Research Council (EPSRC), Swindon, UK. This thesis is dedicated to the memory of Katerina, who had been for long my closest companion. This thesis is dedicated to my parents, Vasileios and Elisavet, who stoically provided unconditional support throughout my studies. Finally, this thesis is dedicated to Lorena, and to the child she is carrying.

Declarations

I declare that the present thesis was composed by myself, and that the work contained herein is my own except where stated otherwise in the text. For the conception and the completion of the work included in Chapters 2 and 3, I collaborated with my first supervisor, Professor Vassili Kolokoltsov. Chapter 2 has been published in the peer-reviewed journal *Games.* For the conception and the completion of the work included in Chapter 4, I collaborated with my second supervisor, Professor Steve Alpern. Chapter 4 has been submitted for publication in the peer-reviewed journal *Operations Research* (at the date of initial submission of this thesis it is under review).

Abstract

Decision making in the context of crime execution and crime prevention can be successfully investigated with the implementation of game-theoretic tools. Evolutionary and mean-field game theory allow for the consideration of a large number of interacting players organized in social and behavioural structures, which typically characterize this context. Alternatively, 'traditional' game-theoretic approaches can be applied for studying the security of an arbitrary network on a two player non-cooperative game. Theoretically underpinned by these instruments, in this thesis we formulate and analyse game-theoretic models of inspection, corruption, counter-terrorism, patrolling, and similarly interpreted paradigms. Our analysis suggests optimal strategies for the involved players, and illustrates the long term behaviour of the introduced systems. Our contribution is towards the explicit formulation and the thorough analysis of real life scenaria involving the security in network structures.

Chapter 1

Introduction

This thesis consists of five chapters, the first chapter being the Introduction, and the last chapter being the Conclusion. The main body consists of Chapters 2, 3, and 4.

In each one of Chapters 2, 3, and 4, a game-theoretic problem is described, formulated, and thoroughly analysed. Although the aforementioned models can be studied independently, they form a whole as they all discuss instances of networks security concentrating primarily on inspection, corruption, counter-terrorism, and patrolling. The network element in our modelling is introduced either through the idea of a large number of interacting players organized in social and behavioural structures, or through the traditional notion of two player games played on graphs. Game theory provides the necessary scientific mechanisms required to investigate these and relative contexts. In principle, game theory studies the strategic interaction (conflict or cooperation) of rational individuals, that is, a game can be thought of as a multi-agent decision problem in a strategic setting. This element of strategic decision making where individuals' actions are interdependent is what distinguishes game theory from decision theory. For a general introduction one can recommend, e.g., Kolokoltsov and Malafeyev [2010], Myerson [2013], Osborne and Rubinstein [1994]. The publication of the *Theory of games and economic behavior* by von Neumann and Morgenstern in 1944, Von Neumann and Morgenstern [2007], is widely recognised as the first formal establishment of the field. Game theory has rapidly advanced ever since, and has been used almost in every branch of social, natural and formal sciences. Some of its prominent representatives were awarded the Nobel prize in economics for their contribution to sub-fields of game theory over the latest years; John Nash, John Harsanyi and Reinhard Selten (1994), Robert Aumann and Thomas Schelling (2005), Leonid Hurwicz, Eric Maskin and Roger Myerson (2007), Alvin Roth and Lloyd Shapley (2012). In what follows, we briefly summarize what each of our introduced game-theoretic models entails. We strictly refer to dynamic game paradigms, that is, to strategic interactions that reoccur over time, so that decision making of the agents at any time influences the evolution of their system's state and thus their future reaction.

In Chapter 2, we extend a standard two-person, non-cooperative, non-zero sum, imperfect inspection game, including a large number of interacting inspectees and a single inspector. Each inspectee adopts one strategy within a finite/infinite bounded set of strategies returning increasingly illegal profits, including compliance with the established rules. The inspectees may periodically update their strategies after randomly inter-comparing the obtained payoffs, setting their collective behaviour subject to evolutionary pressure. Accordingly, the inspector decides, at each update period, the optimum fraction of his/her finite renewable budget to invest on his/her interference with the inspectees' collective effect. To deter the inspectees from violating, the inspector assigns a fine (penalty) to each illegal strategy. We formulate the game mathematically, study its dynamics, and investigate its evolution subject to two key control parameters, the inspection budget and the punishment fine. Introducing a simple linguistic twist, we additionally capture the relative conception of a corruption game.

In Chapter 3, we combine the recently developed mean-field game models of corruption and bot-net defence in cyber-security, along with the evolutionary game approach of inspection and corruption, under an extended scheme including the pressure-resistance game element. We propose a generalised framework for complex interaction in network structures of large number of small players, that includes their individual decision making inside their environment (i.e. the mean-field game component), their binary interactions (i.e. the evolutionary game component), and the pressure of a major player (i.e. the pressure-resistance game component). To perform concrete calculations with this overall complicated model, we suggest working, in turn, in three asymptotic regimes; we assume fast execution of personal decisions, weak binary interactions, and small discounting in time. We attempt to provide a link between the stationary and the time-dependent mean-field game consistency problem.

In Chapter 4, we consider a variation of the recently introduced network patrolling game, where an Attacker carries out an attack on a node of his/her choice for a given number m of consecutive time periods. The critical parameter m indicates the difficulty of the attack, or infiltration, at a given node. To thwart such an attack, the Patroller adopts a walk on the network, aiming to be at the attacked node during one of the attack periods. If this occurs, the attack is intercepted and the Patroller wins the game; otherwise the Attacker wins. To model the important alternative where the Patroller can be identified when he/she is at the Attacker's node, we allow the Attacker to initiate the attack after waiting for a chosen number d of consecutive periods during which the Patroller has been away. Thus, we introduce the term Uniformed Patroller (alternatively we can use the term noisy in contrast to silent) to denote this new information structure. We solve this new version of the network patrolling game, that is more favourable to the Attacker, for various networks: star, line, circle and a mixture. We restrict the Patroller to Markovian strategies, which cover the whole network.

Chapter 2

Evolutionary Inspection and Corruption Games

2.1 Introduction

An inspection game consists of a game-theoretic framework, modelling the non-cooperative interaction between two strategic parties, called the inspector and the inspectee; see, e.g., Avenhaus, von Stengel and Zamir [2002], Avenhaus and Canty [2012], Hohzaki [2013] for a general survey. The inspector aims to verify that certain regulations, imposed by the benevolent principal he/she is acting for, are not violated by the inspectee. On the contrary, the inspectee has a selfish incentive to disobey the established regulations, risking the enforcement of a punishment fine in the case of detection. The introduced punishment mechanism is a key element in the analysis of inspection games, since deterrence is generally considered to be the inspector's highest priority. Typically, the inspector has limited means of inspection at his/her disposal, so that his/her detection efficiency can only be partial.

The central objective of inspection games is to develop an effective inspection policy for the inspector to adopt, given that the inspectee acts according to a strategic plan. Within the last five decades, inspection games have been applied in the game-theoretic analysis of a wide range of issues, mainly in arms control and nuclear non-proliferation, but also in the accounting and auditing of accounts, in tax inspections, environmental protection, crime control, passenger ticket control, stock-keeping and many others; see, e.g., Alferov, Malafeyev and Maltseva [2015], Avenhaus, von Stengel and Zamir [2002], Avenhaus [2004], Deutsch et al. [2013] and the references therein. Though when initially introduced, inspection games appeared almost exclusively as two-person, zero-sum, non-cooperative games, the need to depict more realistic, and therefore of increased complexity, scenaria gradually shifted attention towards N-person and non-zero-sum games.

Dresher's two-person, zero-sum, perfect recall, recursive inspection game, Dresher [1962], is widely recognized as the first formal approach in the field. In his model, Dresher considered n periods of time available for an inspectee to commit, or not, a unique violation, and $m \leq n$ one-period lasting inspections available for the inspector to investigate the inspectee's abidance by the rules, assuming that a violator can be detected only if he/she is caught (inspected) in the act. This work initiated the application of inspection games to arms control and disarmament; see, e.g., Avenhaus et al. [1996] and the references therein. Maschler [1966] generalized this archetypal model, introduced the equivalent non-zero-sum game and, most importantly, adopted from economics the notion of inspector leadership, showing (among others) that the inspector's option to pre-announce and commit to a mixed inspection strategy actually increases his/her expected payoff.

Thomas and Nisgav [1976] used a similar framework to investigate the problem of a patroller aiming to inhibit a smuggler's illegal activity. In their so-called customs-smuggler game, customs patrol, using a speedboat, in order to detect a smuggler's motorboat attempting to ship contraband through a strait. They introduced the possibility of more than one patrolling boats, namely the possibility of two or more inspectors, potentially not identical, and suggested the use of linear programming methods for the solution of those scenario. Baston and Bostock [1991] provided a closed-form solution for the case of two patrolling boats, and discussed the withdrawal of the perfect-capture assumption, stating that detection is ensured whenever violation and inspection take place at the same period. Garnaev [1994] provided a closed-form solution for the case of three patrolling boats.

Von Stengel [1991] introduced a third parameter in Dresher's game, allowing multiple violations, but proving that the inspector's optimal strategy is independent of the maximum number of the inspectee's intended violations. He studied another variation, optimizing the detection time of a unique violation that is detected at the following inspection, given that inspection does not currently take place. On a later version, von Stengel [2016] additionally considered different rewards for the inspectee's successfully committed violations, extending as well Maschler's inspector leadership version under the multiple intended violations assumption. Ferguson and Melolidakis [1998], motivated by Sakaguchi [1977], treated a similar three-parameter, perfect-capture, sequential game, where: (i) the inspectee has the option to 'legally' violate at an additional cost; (ii) a detected violation does not terminate the game; (iii) every non-inspected violation is disclosed to the inspector at the following stage.

Non-zero-sum inspection games were already discussed at an early stage by Maschler [1966, 1967], but were mainly developed after the 1980s, in the context of the nuclear non-proliferation treaty (NPT). The prefect-capture assumption was partly abandoned, and errors of Type 1 (false alarm) and Type 2 (undetected violation given that inspection takes place) were introduced to formulate the so-called imperfect inspection games. Avenhaus and von Stengel [1992] solved Dresher's perfect-capture, sequential game, assuming non-zero-sum payoffs. Canty, Rothenstein and Avenhaus [2001] solved an imperfect, non-sequential game, assuming that players ignore any information they collect during their interaction, where an illegal action must be detected within a critical timespan before its effect is irreversible. They discussed the sequential equivalent, as well. Rothenstein and Zamir [2002] included the elements of imperfect inspection and timely detection in the context of environmental control, extending Diamond's models for a single inspection, Diamond [1982].

Avenhaus and Kilgour [2004] introduced a non-zero-sum, imperfect (Type 2 error) inspection game, where a single inspector can continuously distribute his/her effort-resources between two non-interacting inspectees, exempted from the simplistic dilemma whether to inspect or not. They related the inspector's detection efficiency with the inspection effort through a non-linear detection function and derived results for the inspector's optimum strategy subject to its convexity. Hohzaki [2007] moved two steps forward, considering a similar $n + 1$ players inspection game, where the single inspection authority not only intends to optimally distribute his effort among n inspectee countries, but also among l_k facilities within each inspectee country k. Hohzaki presents a method of identifying a Nash equilibrium for the game and discusses several properties of the players' optimal strategies.

In the special case when the inspector becomes himself/herself the individual under investigation, namely when the philosophical question "Who will guard the guardians?" eventually arises (first stated in the work of the Roman satirist Juvenal, *Satire VI*, Green [2004]), see Hurwicz [2008], the exact same framework can be used for modelling corruption. In the so-called corruption games, a benevolent principal aims to ensure that his/her non-benevolent affiliate does not intentionally fail his/her duty; see, e.g., Aidt [2003], Jain [2001], Kolokoltsov and Malafeyev [2015], Malafeyev, Redinskikh and Alferov [2014] and the references therein for a general survey.

For example, in the tax audit/inspection regime, the tax inspectors employed by the respective competent authority are often open to bribery from the tax payers in order not to report detected tax evasions. Generally speaking, when we switch from inspection to corruption games, the competing pair of an inspector versus an inspectee is replaced by the pair of a benevolent principal versus a non-benevolent employee, but the framework of analysis that is used for the first one can almost identically be applied for the second one, as well.

Lambert-Mogiliansky, Majumdar and Radner [2008] developed a dynamic game where various private investors anticipate the processing of their applications by an ordered number of low level bureaucrats in order to ensure specific privileges; such an application is approved only if every bureaucrat is bribed. Nikolaev [2014] introduced a game theoretic study of corruption with a hierarchical structure, where inspectors of different levels audit the inspectors of the lower level and report (potentially false reports) to the inspectors of the higher level; the inspector of the highest level is assumed to be honest. In the context of ecosystem management and biodiversity conservation, Lee et al. [2015] studied an evolutionary game, where they analyse illegal logging with respect to the corruption of forest rule enforcers, while in the context of politics and governance, Giovannoni and Seidmann [2014] investigated how power may affect the government dynamics of simple models of a dynamic democracy, assuming that "*power corrupts and absolute power corrupts absolutely*" (the famous quote of the British politician Lord Acton).

In this chapter we focus on the study of inspection (and corruption) games from an evolutionary perspective, aimed at the analysis of the class of games with a large number of inspectees. However, we highlight that our setting should be distinctly separated from the general setting of the standard evolutionary game theory. We emphasize the networking aspects of these games by allowing the inspectees to communicate with each other and update their strategies purely on account of their interactions. This way, we depict the real-life scenario of partially-informed, optimizing, interacting, indistinguishable agents. For the same purpose, we set the inspectees to choose from different levels of illegal behaviour. Additionally, we introduce the inspector's budget as a distinct parameter of the game, and we measure his/her interference with the interacting inspectees with respect to this. We examine carefully the critical effect of the punishment fine on the evolution of the game. In fact, we attempt to get quantitative insights into the interplay of these key parameters and analyse respectively the dynamics of the game.

For a real-world implementation of our game, one can think of tax inspections. Tax payers are ordinary citizens who interact on a daily basis exchanging information on various issues. Arguably, in their vast majority, if not universally, tax payers have a selfish incentive towards tax evasion. Depending on the degree of confidence they have in their fellow citizens, on a pairwise level, they discuss their methods, the extent to which they evade taxes and their obtained payoffs. As experience suggests, interacting agents, and therefore the tax payers as well, imitate the more profitable strategies. The tax inspector (say the chief of the tax department) is in charge of fighting tax evasion. Having to deal with many tax payers, primarily he/she aims to confront their collective effect rather than each one individually. Then, provided with a bounded budget from a superior authority (say the finance ministry), the tax inspector aims to manage this, along with his/her punishment policy, so that he/she maximizes his/her utility (namely the payoff of the tax department).

Though we restrict ourselves to the use of inspection game terminology, our model also intends to capture the relevant class of corruption games as those are introduced above. Indicatively, we aim to investigate the dynamics of the interaction between a large group of corrupted bureaucrats and their incorruptible superior, again from an evolutionary perspective. In accordance with our earlier approach, the bureaucrats discuss in pairs their bribes and copy the more efficient strategies, while their incorruptible superior aims to choose attractive wages to discourage bribery, to invest in means of detecting fraudulent behaviour and to adopt a suitable punishment policy. Evidently, the two game settings are fully analogous, and despite the linguistic twist of inspection to corruption, they can be formulated in an identical way.

We organize Chapter 2 as follows. In Section 2.2, we discuss the standard setting of a two-player, non-cooperative, non-zero-sum inspection game, and we introduce what we call the conventional inspection game. In Sections 2.3 and 2.4, we present our generalization; we extend the two-player inspection game considering a large population of indistinguishable inspectees, interacting against a single inspector, we formulate our model for a discrete and a continuous strategy setting respectively, and we demonstrate our analysis of the system's dynamics. In Section 2.5, we include a game-theoretic interpretation of our fixed points analysis.

2.2 Standard Inspection Game

A standard inspection game describes the strictly competitive interaction between an inspectee and an inspector, whose interests in principle contradict. The inspectee, having to obey certain rules imposed by the inspector, either chooses indeed to comply with the established rules, obtaining a legal profit, $r > 0$, or to violate them, targeting at an additional illegal profit, $\ell > 0$, but undertaking additionally the risk of being detected and, consequently, having to pay the corresponding punishment fine, $f > 0$. Likewise, the inspector chooses either to inspect at some given inspection cost, $c > 0$, in order to detect any occurring violation, ward off the loss from the violator's illegal profit and receive the debited fine, or not to inspect, avoiding the cost of inspection, but risking the occurrence of a non-detected violation.

In this two-player game-theoretic setting, both parties are considered to be rational optimizers who decide their strategies independently of each other, without observing or being informed about their competitor's behaviour. Namely, the game under discussion is a non-cooperative one. The following 2×2 normal-form table (Table 2.1) illustrates the framework we described above, where the inspectee is the row player and the inspector is the column player. Left and right cells' entries correspond to the inspectee's and the inspector's payoffs respectively. Notice that, in general, the game is formulated as a non-zero-sum one.

	Inspect	**Not Inspect**
Violate	$r - f,\ -c + f$	$r + \ell,\ -\ell$
Comply	$r,\ -c$	$r,\ 0$

Table 2.1: Two-player perfect inspection game.

Table 2.1 illustrates the so-called perfect inspection game, in the sense that inspection always coincides with detection (i.e. given that a violator is inspected, the inspector will detect his/her violation with probability one). However, this is an obviously naive approach, since in practice, numerous factors deteriorate the inspector's efficiency and potentially obstruct detection (recall the errors of Type 1 and Type 2 we mentioned above). Consequently, the need to introduce a game parameter determining the inspection's efficiency naturally arises.

In this more general setting, the critical parameter $\lambda \in [0, 1]$ is introduced to measure the conditional probability with which a violation is detected given that the inspector conducts an inspection. Alternatively, one can think of λ as a measure of the inspector's detection efficiency. Obviously, for $\lambda = 1$, the ideal scenario of perfect inspection is captured, while, for $\lambda = 0$, detection can be never achieved. The following 2×2 normal-form table (Table 2.2) illustrates the so-called imperfect inspection game.

	Inspect	**Not Inspect**
Violate	$r + \ell - \lambda \cdot (\ell + f),\ -c - \ell + \lambda \cdot (\ell + f)$	$r + \ell,\ -\ell$
Comply	$r,\ -c$	$r,\ 0$

Table 2.2: Two-player imperfect inspection game.

The key feature of the discussed game setting is that under specific conditions, it describes a two-player competitive interaction without any pure strategy Nash equilibria. Starting from the natural assumption that the inspector, in principle, would like the inspectee to comply with his/her rules, and that ideally he/she would prefer to ensure compliance without having to inspect, the game obtains no pure strategy Nash equilibria when both of the following two conditions apply

$$-c - \ell + \lambda \cdot (\ell + f) > -\ell \Rightarrow \lambda \cdot (f + \ell) > c, \tag{2.1}$$

$$r > r + \ell - \lambda \cdot (\ell + f) \Rightarrow \lambda \cdot (f + \ell) > \ell. \tag{2.2}$$

Indicatively, one can verify that the pure strategy profile (V, I), where V stands for Violate and I stands for Inspect, is the unique Nash equilibrium of the imperfect inspection game when only condition (2.1) applies. Accordingly, profile (V, NI), where NI stands for Not Inspect, is the unique pure strategy Nash equilibrium when only condition (2.2) applies. When neither of the two conditions apply, profile (V, NI) is again the unique pure strategy Nash equilibrium. Hence, given that at least one of the above conditions (2.1) and (2.2) does not apply, a pure strategy equilibrium solution always exists.

Back to the no pure strategy Nash equilibria environment, the first condition assumes that when the inspectee is violating, the inspector's expected payoff is higher when he/she chooses to inspect. Accordingly, the second condition assumes that when the inspector is inspecting, the inspectee's expected payoff is higher when he/she chooses to comply (note that this is always true for the perfect inspection game of Table 2.1).

Under these assumptions, regardless of the game's outcome and given the competitor's choice, both players would in turn switch their previously chosen strategies to the alternative ones, in an endlessly repeated switching cycle (see Figure 2.1). This lack of no-regrets pure strategies states that the game contains no pure strategy Nash equilibria. We name it the conventional inspection game.

Comply → Not Inspect → Violate → Inspect → Comply → ...

Figure 2.1: No pure strategy Nash equilibria conventional inspection game.

Typically, a two-player game without any pure strategy Nash equilibria is resolved by having at least one of the players randomising over his/her available pure strategies. In this specific scenario, it can be proven that both players resort to mixed strategies, implying that both inspection and violation take place with non-zero probabilities. In particular, the following theorem proven in Kolokoltsov and Malafeyev [2010] gives the unique mixed strategy Nash equilibrium of the conventional inspection game described above.

Theorem 1. *Let $p \in [0,1]$ be the probability with which the inspectee violates and $q \in [0,1]$ be the probability with which the inspector inspects. The unique mixed strategy Nash equilibrium of the two-player inspection game described in Table 2.2 along with conditions* (2.1) *and* (2.2) *is the mixed strategy profile (p^*, q^*) with:*

$$p^* = \frac{c}{\lambda \cdot (f+\ell)} \quad , \quad q^* = \frac{\ell}{\lambda \cdot (f+\ell)}. \tag{2.3}$$

Proof. See Kolokoltsov and Malafeyev [2010].

□

2.3 Discrete Strategy Setting

Let us now proceed with the natural extension of the two-player game we introduced in the previous section to the real-life scenario of a multi-player problem.

We consider a large population of N indistinguishable, pairwise interacting inspectees exchanging opinions under the pressure of a single inspector. Equivalently in the context of corruption games, one can think of N indistinguishable, pairwise interacting bureaucrats against their incorruptible superior. The game mechanism can be summarized into the following dynamic process.

Initially, the N inspectees decide their strategies individually. They retain their group's initial strategy profile for a certain time span, but beyond that point, on account of the inspector's response to their collective effect, some of the inspectees are eager to update and switch to evidently more profitable strategies. In principle, we assume that an inspectee is an updater with a non-zero probability ω that is characteristic of the inspectees' population.

Indicatively, assume on a periodic basis, and in particular at the beginning of each update period, that an updater discusses his/her payoff with another randomly-chosen inspectee, who is not necessarily an updater himself/herself. If the two interacting inspectees have equal payoffs, then the updater retains his/her strategy. If, however, they have a payoff gap, then the updater is likely to revise his/her strategy, subject to how significant their payoffs' difference is.

Clearly, we do not treat the inspectees as strictly rational optimizers. Instead, we assume that they periodically compare their obtained payoffs in pairs, and they mechanically copy more efficient strategies purely in view of their pairwise interaction and without necessarily being aware of the overall prevailing crime rate or the inspector's response. This assumption is described as the myopic hypothesis, and we introduce it to illustrate the lack of perfect information and the frequently adopted imitating behaviour in various multi-agent social systems. However, as we will see in Section 2.5, ignoring the myopic hypothesis in a strictly game-theoretic context, we can still interpret our results.

Regarding the inspector's response, we no longer consider his/her strategy to be the choice of the inspection frequency (recall the inspector's dilemma in the standard game setting whether to inspect or not). Instead, we take into account the overall effort the inspector devotes to his/her inspection activity. In particular, we identify this generic term as the fraction of the available budget that he/she invests on his/her objective, making the assumption that the inspection budget controls every factor related with his/her detection effectiveness (e.g., the inspection frequency, the no-detection probability, the false alarms, etc.).

At each update event, we assume that the inspector is limited to the same finite, renewable available budget B. Without experiencing any policy-adjusting costs, he/she aims at maximizing his/her payoff against each different distribution of the inspectees' strategies at the least possible cost. Additionally, we assume that at each time point, he/she is perfectly informed about the inspectees' collective behaviour. Therefore, we treat the inspector as a rational, straightforward, payoff maximizing player. This suggestion is described as the best response principle.

Under this game-theoretic framework, the distribution of violating the established regulation in the population of the inspectees is subject to evolutionary pressure over time. As a result, the term evolutionary is introduced to describe the inspection (corruption) game. It turns out that the more efficient strategies gradually become dominant through imitation.

2.3.1 Analysis

We initiate our analysis by assuming that the inspectees choose their strategies within a finite bounded set of strategies $S = \{0, 1, ..., d\}$, generating increasingly illegal profits. Their group's state space is then the set of sequences of $d+1$ non-negative integers $n = (n_0, \dots, n_d)$, n_i denoting the occupation frequency of strategy $i \in S$. Equivalently, it is the set of sequences of the corresponding $d+1$ relative occupation frequencies $x = (x_0, \dots, x_d)$, where $x_i = n_i/N$.

We consider a constant number of inspectees, namely we have $N = n_0 + \dots + n_d$ for every group's state n. Provided that the population size N is sufficiently large (formally the following is valid for $N \to \infty$ through the law of large numbers), we approximate the relative occupation frequencies x_i with $\rho_i \in [0, 1]$, denoting the probabilities with which the strategies $i \in S$ are adopted. To each strategy i we assign an illegal profit ℓ_i, $\ell_0 = 0$ characterizing the compliers, and a strictly increasing punishment fine $f_i = f(\ell_i)$, with $f_0 = 0$. We assume that $|\ell_{i+1} - \ell_i|$ is constant, namely that ℓ_i's form a one-dimensional regular lattice.

As explained, the inspector has to deal with an evolving crime distribution in the population of the inspectees, that is $\mathrm{p} = \mathrm{p}(t) = (\rho_i)(t)$. We thus define the set of probability vectors Σ_{d+1}, such that:

$$\Sigma_{d+1} = \Big\{ \mathrm{p}(t) = (\rho_0, \dots, \rho_d)(t) \in \mathrm{R}_+^{\mathrm{d}+1} : \sum \rho_i(t) = 1 \Big\}. \tag{2.4}$$

We introduce the inner product notation to define the group's expected (average) illegal profit by $\bar{\ell} = \bar{\ell}(\mathrm{p}) = \langle \ell, \mathrm{p} \rangle$. Respectively, we define the group's expected (average) punishment fine by $\bar{f} = \bar{f}(\mathrm{p}) = \langle f, \mathrm{p} \rangle$. We also define the inspector's invested budget against crime distribution p by $b(\cdot) \in [0, B]$ and the inspector's efficiency by $G(b)$. The last function measures the probability with which a violating inspectee is detected given that the inspector invests budget b.

To depict a plausible scenario, we assume that perfect efficiency cannot be achieved within the inspector's finite available inspection budget B (namely, the detection probability is strictly smaller than one, $G(B) < 1$).

Assumption 1. *The inspector's efficiency, $G : [0, \infty) \mapsto [0, 1)$, is a twice continuously differentiable, strictly increasing, strictly concave function, satisfying:*

$$G'' < 0, \quad G' > 0, \quad \lim_{x \to 0} G'(x) = \infty, \quad G(0) = 0, \quad \lim_{x \to \infty} G(x) = 1.$$

An inspectee who plays strategy $i \in S$, either escapes undetected with probability $1 - G(b)$ and obtains an illegal profit ℓ_i, or gets detected with probability $G(b)$ and is charged with a fine f_i. Additionally, every inspectee receives a legal income r, regardless of his/her strategy being legal or illegal.

Therefore, to an inspectee playing strategy i, against the inspector investing budget b, we assign the following inspectee's payoff function:

$$\Pi_i(b) = r + \big(1 - G(b)\big) \cdot \ell_i - G(b) \cdot f_i. \tag{2.5}$$

Accordingly, we need to introduce a payoff function for the inspector investing budget b against a crime distribution p. Recall that the inspector is playing against a large population of indistinguishable, interacting agents and intends to suppress their collective illegal behaviour. That being the case, for his/her macroscopic assessment, the larger the group is, the less considerable the absolute values corresponding to a single agent (i.e., r, ℓ_i, f_i) are. To depict this inspector's subjective evaluation, we introduce the inspector's payoff function as follows:

$$\Pi_I(b, \mathrm{p}, N) = -b + N \cdot G(b) \cdot \bar{f}_t \cdot \frac{\kappa}{N} - N \cdot \big(1 - G(b)\big) \cdot \bar{\ell}_t \cdot \frac{\kappa}{N}, \tag{2.6}$$

where κ is a positive scaling constant and $\bar{\ell}_t$, respectively $\bar{f}_t$, denotes the expected (average) illegal profit, respectively the expected (average) punishment fine, at time t. Without loss of generality, we can set $\kappa = 1$. Note that the inspector's payoff always obtains a finite value, including the limit $N \to \infty$.

As already mentioned, an inspectee (updater) revises his/her strategy with a switching probability depending on his/her payoff's difference with another randomly-chosen individual's payoff, with whom he/she discusses outcomes. Then, for an updater playing strategy $i \in S$ and exchanging information with an inspectee playing strategy $j \in S$, we define this switching probability by $s_{ij} \cdot \Delta t$, for a timespan Δt, where:

$$s_{ij} = \begin{cases} \beta \cdot \big(\Pi_j(b) - \Pi_i(b)\big) & , \quad \text{if } \Pi_j(b) > \Pi_i(b) \\ 0 & , \quad \text{if } \Pi_j(b) \leq \Pi_i(b) \end{cases} \tag{2.7}$$

and $\beta > 0$ is an appropriately-scaled normalization parameter.

This transition can be summarised in the following dynamic process; in every period following an update event, the number of inspectees playing strategy i is equal to the corresponding sub-population in the previous period, plus the number of inspectees having previously played strategies $j \neq i$ and switching now into strategy i, minus the number of inspectees having previously played strategy i and switching now into strategies $j \neq i$.

Hence, we derive the following iteration formula:

$$\rho_i(t+\Delta t) = \rho_i(t) + \omega \cdot \rho_i(t) \cdot \sum_{j=0}^{d} \big(\rho_j(t) \cdot s_{ij}\big) \cdot \Delta t - \omega \cdot \rho_i(t) \cdot \sum_{j=0}^{d} \big(\rho_j(t) \cdot s_{ji}\big) \cdot \Delta t \tag{2.8}$$

which can be suitably reformulated, taking the limit as $\Delta t \to 0$, into an equation resembling the well-known replicator equation (see, e.g., Zeeman [1980]):

$$\dot{\rho}_i(t) = \omega \cdot \beta \cdot \rho_i(t) \cdot \Big(\Pi_i(b) - \sum_{j \in S} \rho_j \cdot \Pi_j(b)\Big). \tag{2.9}$$

Remark 1. *We have used here a heuristic technique to derive equation* (2.9)*, bearing in mind that we consider a significantly large group of interacting individuals (formally valid for the limiting case of an infinitely large population). A rigorous derivation* *with the use of the law of large numbers for interacting Markov chains can be found for example in Kolokoltsov [2012] or Kolokoltsov [2014].*

In agreement with our game setting (in particular with the myopic hypothesis), equation (2.9) is not a best-response dynamic. However, it turns out that successful strategies, yielding payoffs higher than the group's average payoff, are subject to evolutionary pressure. This interesting finding of our setting, which is put forward in the above replicator equation, simply states that although the inspectees are not considered to be strictly rational maximizers (but instead myopic optimizers), successful strategies propagate into their population through the imitation procedure. This characteristic classifies equation (2.9) into the class of the payoff monotonic game dynamics, see, e.g., Hofbauer and Sigmund [1998].

Before proceeding any further, it is important to state that our setting is quite different from the general setting of standard evolutionary game theory. Unlike standard evolutionary games, in our approach there are no small games of a fixed number of players through which successful strategies evolve. On the contrary, at each step and throughout the whole procedure, there is only one $N+1$ players game taking place (see also the analysis in Section 2.5).

In regard to the inspector's interference with the interacting inspectees, the best response principle states that at each time step, against the crime distribution he/she confronts, the inspector aims to maximize his/her instantaneous payoff with respect to his/her available budget:

$$\max_{b\in[0,B]}\big\{-b+G(b)\cdot\bar{f}_t-\big(1-G(b)\big)\cdot\bar{\ell}_t\big\}. \tag{2.10}$$

On the one hand, the inspector chooses his/her fine policy strategically in order to manipulate the evolution of the future crime distribution. On the other hand, at each update period, he/she has at his/her disposal the same finite renewable budget B, while he/she is not charged with any policy adjusting costs. Namely, the inspector has a period-long planning horizon regarding his/her financial policy, and he/she instantaneously chooses at each step his/her response b that maximizes his/her payoff (2.6) against the prevailing crime distribution.

Let us define the inspector's best response (optimum employable budget), maximizing his/her payoff (2.6) against the prevailing crime distribution, by:

$$\hat{b}(\cdot):=\operatorname*{argmax}_{b\in[0,B]}\big\{-b+G(b)\cdot\big(\bar{f}_t+\bar{\ell}_t\big)-\bar{\ell}_t\big\}. \tag{2.11}$$

Having analytically discussed the inspectees' and the inspector's individual dynamic characteristics, we can now combine them and obtain a clear view of the system's dynamic behaviour as a whole.

In particular, we substitute the inspector's best response (optimum employable budget) $\hat{b}(\cdot)$ into the system of ordinary differential equations (ODEs) (2.9), and we obtain the corresponding system governing the dynamic evolution of the non-cooperative game described above:

$$\dot{\rho}_i(t) = \omega \cdot \beta \cdot \rho_i(t) \cdot \big(\ell_i - (\ell_i + f_i) \cdot G(\hat{b}) + G(\hat{b}) \cdot (\bar{\ell} + \bar{f}) - \bar{\ell}\big). \tag{2.12}$$

Recall that through the system (2.12) we aim to investigate the evolution of illegal behaviour within a large group of interacting, myopically-maximizing, indistinguishable inspectees (bureaucrats) under the pressure of a single rationally-maximizing inspector (incorruptible superior).

Without loss of generality, we can set the normalization parameter $\beta = 1$. Let us also introduce the following auxiliary notation:

$$K_i(\mathrm{p}, \hat{b}) = \ell_i - (\, \ell_i + f_i\,) \cdot G(\hat{b}) + G(\hat{b}) \cdot (\bar{\ell} + \bar{f}) - \bar{\ell}. \tag{2.13}$$

Proposition 1. *A probability vector* $\mathrm{p}(t) \in \Sigma_{d+1}$ *is a singular point of* (2.12)*, namely it satisfies the system of equations:*

$$\omega \cdot \rho_i(t) \cdot K_i(\mathrm{p}, \hat{b}) = 0, \tag{2.14}$$

if and only if there exists a subset $I \subset S$*, such that* $\rho_i(t) = 0$ *for* $i \in I$*, and* $K_i(\mathrm{p}, \hat{b}) = 0$ *for* $i \notin I$*.*

Proof. For any subset $I \subset S$ such that $\rho_i(t) = 0$, $i \in I$, system (2.12) reduces to the same system, but only with the coordinates $i \notin I$ (notice that I must be a proper subset of S). Then, for the fixed point condition to be satisfied, we must have $K_j(\mathrm{p}, \hat{b}) = 0$ for every $j \in S \backslash I$.

□

The determination of the fixed points defined in Proposition 1, as well as their stability analysis (namely the deterministic evolution of the game), clearly depend on the explicit form of the auxiliary function K_i. One can distinguish, then, two control elements that appear in K_i and thus govern the dynamics of the game; the functional control $f(\cdot)$ and the control parameter B. We have set the fine $f(\cdot)$ to be a strictly increasing function, and we further consider three eventualities regarding its convexity; (i) linear; (ii) convex; (iii) concave. To each of the above three versions we assign a different inspector's punishment profile.

Indicatively, we claim that a convex fine function reveals an inspector who is lenient against relatively low collective violation, but rapidly jumps to stricter policies when coming up against increasing collective violation. On the contrary, we claim that a concave fine function reveals an inspector who is already punishing aggressively even for a relatively low collective violation (pre-empting inspector). Finally, we assume that a linear fine function represents the severity of the 'average' inspector. In each case, we vary the constant, increasing, or decreasing gradient of function $f(\cdot)$, respectively for a linear, convex, or concave fine. Accordingly, we vary the size of the finite available budget B.

The different settings we establish with these control parameter variations, and therefore, the corresponding dynamics we obtain in each occasion, have clear practical interpretation providing useful insight into applications. For example, the fine function $f(\cdot)$ can be, and usually is, defined by the inspector himself (think of the different fine policies when dealing with tax evasion), while the level of budget B is decided from the benevolent principal by whom the inspector is employed.

Contrariwise, the detection efficiency $G(\cdot)$ is not regarded as an additional control since it characterizes the inspector's behaviour endogenously. However, say the inspector has an excess of budget, which he/she could invest in improving his/her expertise (e.g., the technical know-how). This is related (indirectly) with his/her efficiency, and thus could partially improve $G(\cdot)$. We do not engage with this scenario.

2.3.2 Linear Fine

Equivalently to (2.11), for a linear fine $f_i = \sigma \cdot \ell_i$, $\sigma \in \mathbb{R}^+$, the inspector's best response (optimum employable budget) can be written as:

$$\hat{b}(\bar{\ell}) = \min\Big[B, (G^{'})^{-1}\Big(\frac{1}{\sigma \cdot \bar{\ell}_t + \bar{\ell}_t}\Big)\Big]. \tag{2.15}$$

We conclude from (2.15) that we cannot have $\hat{b}(\bar{\ell}) > B$ for every $\bar{\ell} \in [0, \ell_d]$, since at least for the case when $\bar{\ell} = 0$, it is $\hat{b}(0) = 0$. However, depending on the size of the available budget B, we may have $B > \hat{b}(\bar{\ell})$ for every $\bar{\ell} \in [0, \ell_d]$.

Then, it is reasonable to introduce the following notation:

$$\ell_c := \min\big\{\ell : \hat{b}(\ell) = B\big\}, \tag{2.16}$$

where ℓ_c is not necessarily deliverable, i.e., ℓ_c may not belong to $[0, \ell_d]$. One should think of this critical value ℓ_c as a measure of the adequacy of the inspector's available budget B, namely as the 'strength' of his/her available budget. Obviously, if $\ell_c \leq \ell_d$, the inspector benefits from exhausting all his/her available budget when dealing with a collective violation $\bar{\ell} \in [\ell_c, \ell_d]$, while, if $\ell_c > \ell_d$, the inspector never needs to exhaust B in order to achieve an optimum response.

Theorem 2. *Let Assumption 1 hold. The $d+1$ unit vectors $\mathrm{p}_i = (\delta_{ij})$, $i, j \in S$, lying on the vertices of the d-simplex, are fixed points of* (2.12)*. Moreover:*

1. *If $G(\hat{b}(\ell_d)) > \frac{1}{1+\sigma}$, there is additionally a unique hyperplane of fixed points,*

$$\Theta = \Big\{ \mathrm{p}_\theta \in \Sigma_{d+1} \mid \exists!\, \bar{\ell} \in (0, \min[\ell_c, \ell_d]) : \langle \ell, \mathrm{p}_\theta \rangle = \bar{\ell} \ \ \& \ \ G(\hat{b}(\bar{\ell})) = \frac{1}{1+\sigma} \Big\},$$

2. *If $G(\hat{b}(\ell_d)) = \frac{1}{1+\sigma}$ and $\ell_d > \ell_c$, there are additionally infinitely many hyperplanes of fixed points,*

$$\Phi = \Big\{ \mathrm{p}_\phi \in \Sigma_{d+1} \mid \forall\, \bar{\ell} \in [\ell_c, \ell_d] : \langle \ell, \mathrm{p}_\phi \rangle = \bar{\ell} \ \ \& \ \ G(\hat{b}(\bar{\ell})) = \frac{1}{1+\sigma} \Big\}.$$

Proof. In any case, the unit probability vectors $\mathrm{p}_i = (\delta_{ij})$, $i, j \in S$, satisfy system (2.14), since by the definition of p_i, it is $\rho_j = 0$, $\forall\, j \neq i$, whilst it is $\langle \ell, \mathrm{p}_i \rangle = \ell_i$.

The setting we introduce with Assumption 1 ensures that $\hat{b} : [0, \ell_d] \mapsto [0, \hat{b}(\ell_d)]$ is a continuous, non-decreasing, surjective function. In particular, we have that $G(\cdot)$ is strictly increasing in $\hat{b} \in [0, B]$ and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\ell} \in [0, \min[\ell_c, \ell_d]]$.

Hence, the following hold:

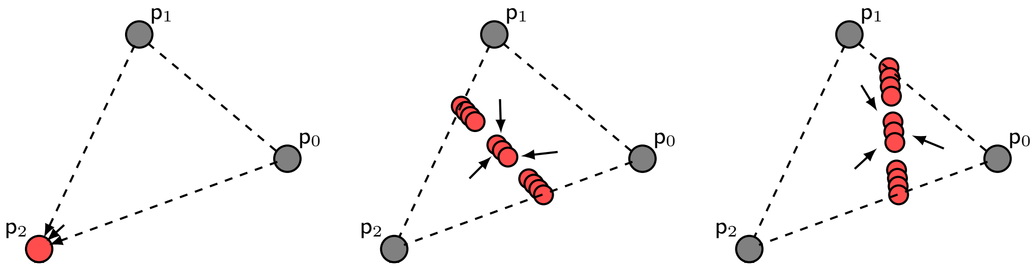
1. When $G(\hat{b}(\ell_d)) > \frac{1}{1+\sigma}$, there is a unique average value $\bar{\ell} \in (0, \min[\ell_c, \ell_d])$ satisfying $G(\hat{b}(\bar{\ell})) = \frac{1}{1+\sigma}$. This unique $\bar{\ell}$ is generated by infinitely many probability vectors, $\mathrm{p}_\theta :\ \langle \ell, \mathrm{p}_\theta \rangle = \bar{\ell}$, forming a hyperplane of vector points satisfying (2.14).

2. When $G(\hat{b}(\ell_d)) = \frac{1}{1+\sigma}$ and $\ell_d > \ell_c$, every average value $\bar{\ell} \in [\ell_c, \ell_d]$ satisfies $G(\hat{b}(\bar{\ell})) = \frac{1}{1+\sigma}$. Each one of these infinitely many $\bar{\ell}$ is generated by infinitely many probability vectors, $\mathrm{p}_\phi : \langle \ell, \mathrm{p}_\phi \rangle = \bar{\ell}$, forming infinitely many hyperplanes of vector points satisfying (2.14).

□

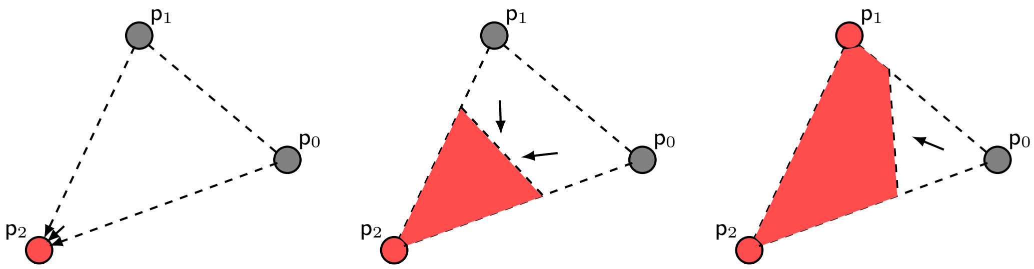
We refer to the vector points p_i as pure strategy fixed points, to emphasize that they correspond to the population's strategy profiles such that every inspectee plays the same strategy $i \in S$. Accordingly, we refer to the vector points p_θ, p_ϕ, as mixed strategy fixed points, to emphasize that they correspond to the population's strategy profiles such that the inspectees are distributed among two or more strategies.

Before proceeding with the general stability results, we present the detailed picture in the simplest case of three available strategies generating increasingly illegal profits including compliance.

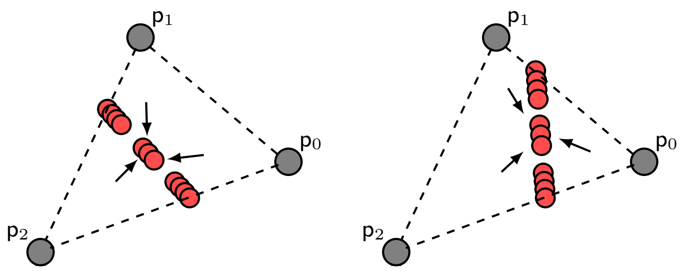


(a) $G(\hat{b}(\ell_2)) \leq \frac{1}{1+\sigma}$ (b) $G(\hat{b}(\ell_1)) < \frac{1}{1+\sigma} < G(\hat{b}(\ell_2))$ (c) $\frac{1}{1+\sigma} < G(\hat{b}(\ell_1)) < G(\hat{b}(\ell_2))$

Figure 2.2: Dynamics for a linear $f(\cdot)$, where $\ell_c \geq \ell_d$. Set of strategies $S = \{0, 1, 2\}$.



(a) $G(\hat{b}(\ell_2)) < \frac{1}{1+\sigma}$ (b) $G(\hat{b}(\ell_1)) < \frac{1}{1+\sigma} = G(\hat{b}(\ell_2))$ (c) $G(\hat{b}(\ell_1)) = \frac{1}{1+\sigma} = G(\hat{b}(\ell_2))$



(d) $G(\hat{b}(\ell_1)) < \frac{1}{1+\sigma} < G(\hat{b}(\ell_2))$ (e) $\frac{1}{1+\sigma} < G(\hat{b}(\ell_1)) < G(\hat{b}(\ell_2))$

Figure 2.3: Dynamics for a linear $f(\cdot)$, where $\ell_d > \ell_c$. Set of strategies $S = \{0, 1, 2\}$.

Figure 2.2 represents a budget B such that the inspector never exhausts it. For a relatively low B, or for an overly lenient $f(\cdot)$ (see Figures 2.2(a), 2.3(a)), the pure strategy fixed point p_2 is asymptotically stable. Increasing though B or, accordingly, toughening up the fine policy $f(\cdot)$ (see Figure 2.2(b)), a hyperplane of asymptotically stable mixed strategy fixed points appears. Depending on the critical parameter ℓ_c, we may have infinitely many hyperplanes of asymptotically stable mixed strategy fixed points (see Figures 2.3(b), 2.3(d)). Finally, keeping B constant, the more we increase the slope of $f(\cdot)$, the closer this(ese) hyperplane(s) moves towards compliance (see Figures 2.2(c), 2.3(c), and 2.3(e)).

We generalize these results into the following theorem.

Theorem 3. *Let Assumption 1 hold. Consider the fixed points given by Theorem 2. Then:*

1. *the pure strategy fixed point* p_0 *is a source, thus unstable;*
2. *the pure strategy fixed points* $\mathrm{p}_j \notin \Theta, \Phi$, $j \in S : j \neq 0, d$, *are saddles, thus unstable;*
3. *the pure strategy fixed point* p_d *is asymptotically stable when* $\Phi = \Theta = \emptyset$*; otherwise, it is a source, thus unstable;*
4. *the mixed strategy fixed points* p_θ, p_ϕ *are asymptotically stable.*

Proof. See Section 2.6.

□

2.3.3 Convex/Concave Fine

Let us introduce the auxiliary variable $\xi_i = \ell_i + f_i$. As we did before, using the inner product notation, we define the corresponding group's expected (average) value by $\bar{\xi} = \bar{\xi}(\mathrm{p}) = \langle \xi, \mathrm{p} \rangle$, where $\bar{\xi} = \bar{\ell} + \bar{f}$. Then, equivalently to expression (2.11), or (2.15), the inspector's best response (optimum employable budget) can be written as:

$$\hat{b}(\bar{\xi}) = \min\Big[B, (G')^{-1}\Big(\frac{1}{\bar{\xi}_t}\Big)\Big]. \tag{2.17}$$

For every $i, j \in S : i < j$, let us introduce as well the parameter:

$$q^{i,j} = \frac{\ell_i - \ell_j}{\ell_i - \ell_j + f_i - f_j}. \tag{2.18}$$

Lemma 1. *For a convex fine,* $q^{i,j}$ *is strictly decreasing in* i *for constant* j *(or vice versa), while for a concave fine,* $q^{i,j}$ *is strictly increasing in* i *for constant* j *(or vice versa). Furthermore, for a convex fine,* $q^{i,j}$ *is strictly decreasing in* i, j *for constant* $(j - i)$*, while for a concave fine,* $q^{i,j}$ *is strictly increasing in* i, j *for constant* $(j - i)$*.*

Theorem 4. *Let Assumption 1 hold. The $d+1$ unit vectors $\mathrm{p}_i = (\delta_{ij})$, $i, j \in S$, lying on the vertices of the d-simplex are fixed points of system* (2.12)*. Moreover, there may be additionally up to $\binom{d+1}{2}$ internal fixed points $\mathrm{p}_{i,j} \in \Sigma_{d+1}$, living on the support of two strategies $i, j \in S : i < j$, uniquely defined for each pair of strategies; these internal fixed points exist given that the following condition applies respectively:*

$$G(\hat{b}(\xi_j)) > q^{i,j} > G(\hat{b}(\xi_i)). \tag{2.19}$$

Proof. In any case, the unit probability vectors $\mathrm{p}_i = (\delta_{ij})$ satisfy system (2.14), since by the definition of p_i, it is $\rho_j = 0$, $\forall\, j \neq i$, whilst it is $\langle \ell, \mathrm{p}_i \rangle = \ell_i$, $\langle f, \mathrm{p}_i \rangle = f_i$.

Consider now a probability vector $\mathrm{p}^* \in \Sigma_{d+1}$ satisfying (2.14), such that $\langle \ell, \mathrm{p}^* \rangle = \ell^*$, $\langle f, \mathrm{p}^* \rangle = f^*$, $\langle \xi, \mathrm{p}^* \rangle = \xi^*$ and $\mathrm{p}^* \neq \mathrm{p}_i$. Then, from Proposition 1, vector p^* should satisfy $K_i(\mathrm{p}^*, \hat{b}) = 0$, $\forall\, i \notin I$, namely the fraction $q^{i,j}$ should be constant $\forall\, i, j \notin I$, and equal to $G(\hat{b}(\xi^*))$.

To satisfy this, according to Lemma 1, the complement set $I' = S \backslash I$ may not contain more than two elements, namely the distributions p^* may live on the support of only two strategies.

For such a distribution $\mathrm{p}^* = \mathrm{p}_{i,j}$, such that $\langle \ell, \mathrm{p}_{i,j} \rangle = \ell_i \cdot \rho_i + \ell_j \cdot \rho_j = \ell_{i,j}$ and $\langle f, \mathrm{p}_{i,j} \rangle = f_i \cdot \rho_i + f_j \cdot \rho_j = f_{i,j}$, $i, j \in S : i < j$, where $\ell_{i,j} + f_{i,j} = \xi_{i,j}$, we get:

$$G(\hat{b}(\xi_{i,j})) = q^{i,j}. \tag{2.20}$$

The setting we introduce with Assumption 1 ensures that $\hat{b} : [0, \xi_d] \mapsto [0, \hat{b}(\xi_d)]$ is a continuous, non-decreasing, surjective function. In particular, we have that $G(\cdot)$ is strictly increasing in $\hat{b} \in [0, B]$ and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\xi} \in [0, \min[\xi_c, \xi_d]]$.

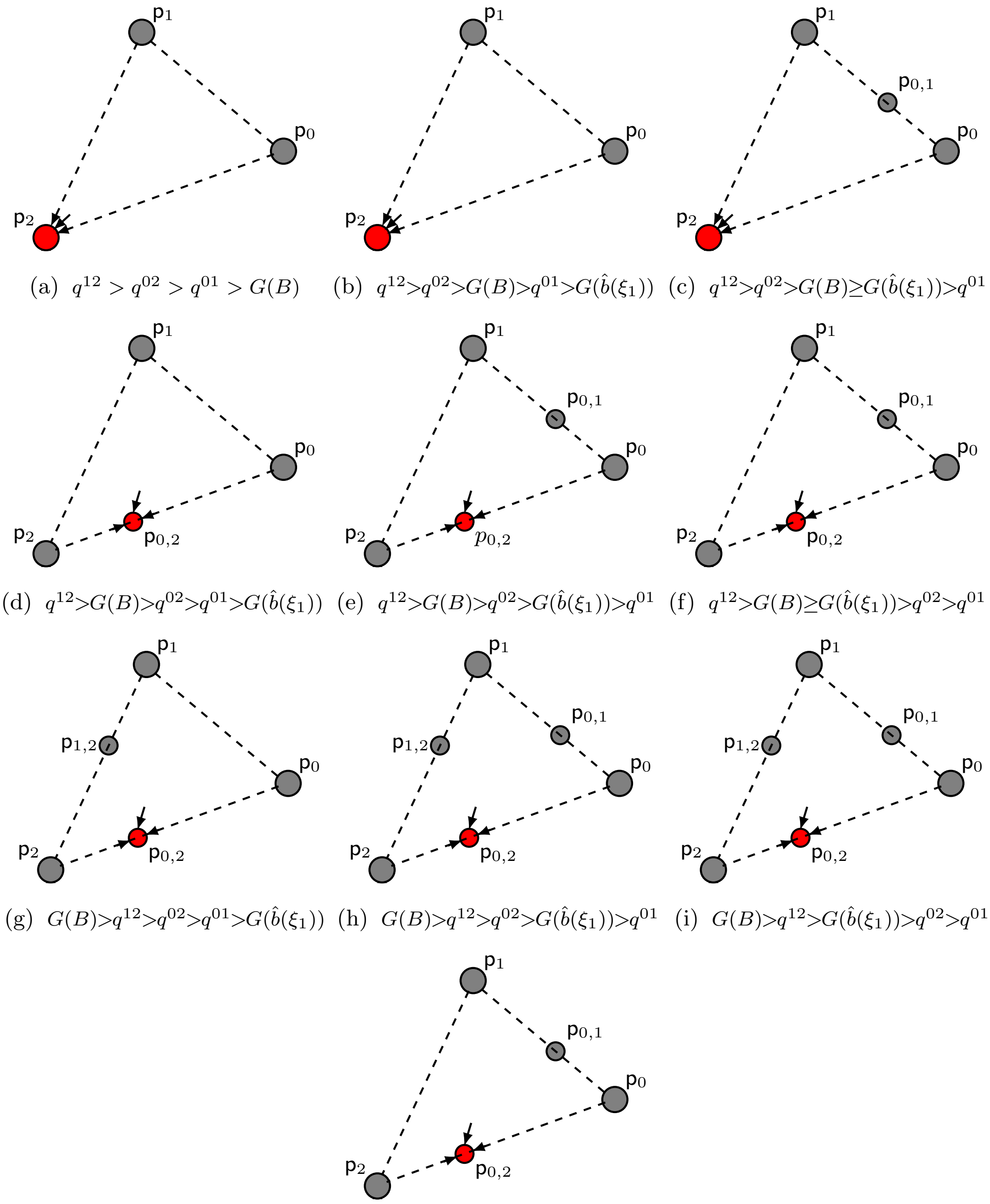
Then, for any $\mathrm{p}_{i,j}$ to exist, namely for (2.20) to hold in each instance, the following condition must hold respectively:

$$G(\hat{b}(\xi_j)) > q^{i,j} > G(\hat{b}(\xi_i)). \tag{2.21}$$

□

We refer to the vector points $\mathrm{p}_{i,j}$ as double strategy fixed points, to emphasize that they correspond to the group's strategy profiles, such that the inspectees are distributed between two available strategies.

Again, we present the detailed picture in the simplest case of three available strategies generating increasingly illegal profits including compliance. Like above, in Figures 2.4 and 2.5, we observe how the interplay of the key control parameters B and $f(\cdot)$ affect the game dynamics. The general pattern is similar to Figures 2.2, 2.3.



(a) $q^{12} > q^{02} > q^{01} > G(B)$ (b) $q^{12}>q^{02}>G(B)>q^{01}>G(\hat{b}(\xi_1))$ (c) $q^{12}>q^{02}>G(B)\geq G(\hat{b}(\xi_1))>q^{01}$

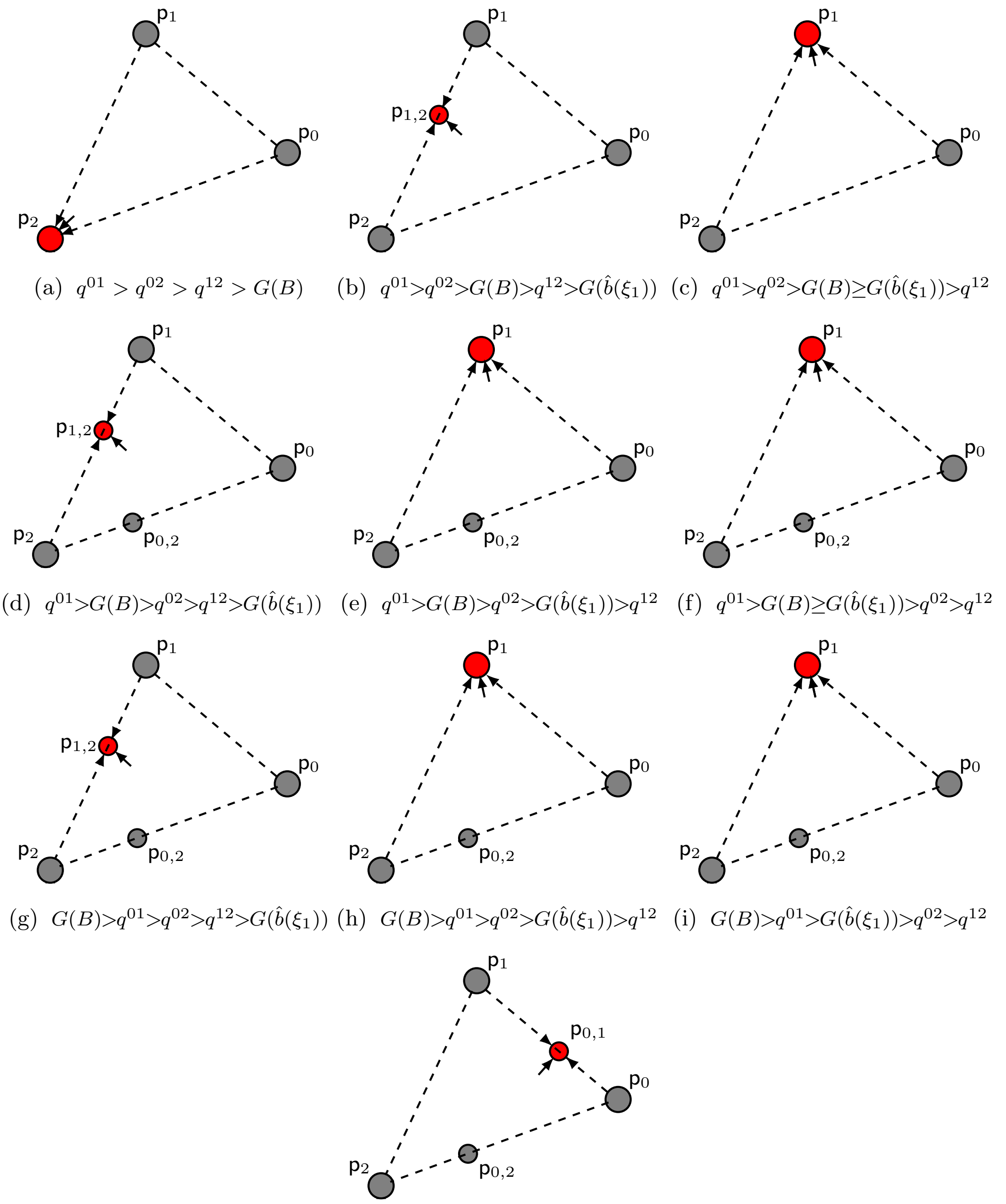
(d) $q^{12}>G(B)>q^{02}>q^{01}>G(\hat{b}(\xi_1))$ (e) $q^{12}>G(B)>q^{02}>G(\hat{b}(\xi_1))>q^{01}$ (f) $q^{12}>G(B)\geq G(\hat{b}(\xi_1))>q^{02}>q^{01}$

(g) $G(B)>q^{12}>q^{02}>q^{01}>G(\hat{b}(\xi_1))$ (h) $G(B)>q^{12}>q^{02}>G(\hat{b}(\xi_1))>q^{01}$ (i) $G(B)>q^{12}>G(\hat{b}(\xi_1))>q^{02}>q^{01}$

(j) $G(B)\geq G(\hat{b}(\xi_1))>q^{12}>q^{02}>q^{01}$

Figure 2.4: Dynamics for the case of concave $f(\ell)$ and three available strategies/$S = \{0, 1, 2\}$

Initially, the pure strategy fixed point p_2 appears to be asymptotically stable (see Figures 2.4(a)–2.4(c), 2.5(a)), but gradually, either increasing B or toughening up $f(\cdot)$, this unique asymptotically stable fixed point shifts towards compliance. However, the shifting in this case takes place through double strategy fixed points, not through hyperplanes of fixed points. For a concave $f(\cdot)$, shifting towards compliance occurs through the fixed point $\mathrm{p}_{0,2}$ living on the support of the two border strategies (see Figures 2.4(d)–2.4(j)), while for a convex $f(\cdot)$, it occurs through the fixed points $\mathrm{p}_{1,2}$, $\mathrm{p}_{0,1}$, living on the support of consecutive strategies (see Figures 2.5(b)–2.5(j)).



(a) $q^{01} > q^{02} > q^{12} > G(B)$ (b) $q^{01}>q^{02}>G(B)>q^{12}>G(\hat{b}(\xi_1))$ (c) $q^{01}>q^{02}>G(B)\geq G(\hat{b}(\xi_1))>q^{12}$

(d) $q^{01}>G(B)>q^{02}>q^{12}>G(\hat{b}(\xi_1))$ (e) $q^{01}>G(B)>q^{02}>G(\hat{b}(\xi_1))>q^{12}$ (f) $q^{01}>G(B)\geq G(\hat{b}(\xi_1))>q^{02}>q^{12}$

(g) $G(B)>q^{01}>q^{02}>q^{12}>G(\hat{b}(\xi_1))$ (h) $G(B)>q^{01}>q^{02}>G(\hat{b}(\xi_1))>q^{12}$ (i) $G(B)>q^{01}>G(\hat{b}(\xi_1))>q^{02}>q^{12}$

(j) $G(B)\geq G(\hat{b}(\xi_1))>q^{01}>q^{02}>q^{12}$

Figure 2.5: Dynamics for the case of convex $f(\ell)$ and three available strategies/$S = \{0, 1, 2\}$

Proposition 2. *Consider the fixed points given by Theorem 4. For a convex fine:*

1. *the set of double strategy fixed points contains at most one fixed point* $\mathrm{p}_{i,i+1}$ *living on the support of two consecutive strategies;*
2. *there is at most one pure strategy fixed point* p_i *satisfying:*

$$q^{i-1,i} > G(\hat{b}(\xi_i)) > q^{i,i+1}. \tag{2.22}$$

Proof. We prove the two statements by contradiction.

1. Assume there are two double strategy fixed points, $\mathrm{p}_{i,i+1}, \mathrm{p}_{j,j+1}$, $i, j \in S$, such that $i < j$, both living on the support of two consecutive strategies. In line with Theorem 4, both of them should satisfy (2.19). However, since we consider a convex fine, then from Lemma 1, it is also $q^{i,i+1} > q^{j,j+1}$, and since it is $j > i$, then from Assumption 1, it is also $G\big(\hat{b}(\xi_j)\big) \geq G(\hat{b}(\xi_{i+1}))$. Overall, we get that:

$$q^{i,i+1} > q^{j,j+1} > G(\hat{b}(\xi_j)) \geq G(\hat{b}(\xi_{i+1})),$$

which contradicts the initial assumption.

2. Assume there are two pure strategy fixed points, p_i, p_j, $i, j \in S$, such that $i \neq 0, i < j$, both satisfying (2.22). However, since we consider a convex fine, then from Lemma 1, it is also $q^{i,i+1} \geq q^{j-1,j}$, and since it is $j > i$, then from Assumption 1, it is also $G(\hat{b}(\xi_j)) \geq G(\hat{b}(\xi_i))$. Overall, we get that:

$$q^{i,i+1} \geq q^{j-1,j} > G(\hat{b}(\xi_j)) \geq G(\hat{b}(\xi_i)),$$

which contradicts the initial assumption.

□

We generalize the findings discussed above on the occasion of Figures 2.4, 2.5, into the following theorem.

Theorem 5. *Let Assumption 1 hold. Consider the fixed points given by Theorem 4. For a concave fine:*

1. *the pure strategy fixed point* p_0 *is a source, thus unstable;*
2. *the pure strategy fixed points* $\mathrm{p}_i, i \neq 0, d,$ *are saddles, thus unstable;*
3. *the double strategy fixed points* $\mathrm{p}_{i,j} \neq \mathrm{p}_{0,d}$ *are saddles, thus unstable;*

4. *the double strategy fixed point* $\mathrm{p}_{0,d}$ *is asymptotically stable;*

5. *the pure strategy fixed point* p_d *is asymptotically stable when* $\mathrm{p}_{0,d}$ *does not exist; otherwise, it is a source, thus unstable.*

For a convex fine:

1. *the pure strategy fixed point* p_0 *is a source, thus unstable;*

2. *the double strategy fixed points* $\mathrm{p}_{i,j}, j \neq i+1$*, are saddles, thus unstable;*

3. * *the double strategy fixed points* $\mathrm{p}_{i,i+1}$ *are asymptotically stable;*

4. * *the pure strategy fixed points* $\mathrm{p}_i, i \neq 0, d$*, are saddles, thus unstable;*

5. * *the pure strategy fixed point* p_d *is a source, thus unstable.*

* *When no double strategy fixed point* $\mathrm{p}_{i,i+1}$ *living on the support of two consecutive strategies exists, the pure strategy fixed point satisfying* (2.22) *is asymptotically stable.*

Proof. See Section 2.6.

□

2.4 Continuous Strategy Setting

The discrete strategy setting is our first approach towards introducing multiple levels of violation available for the inspectees. It is an easier framework to work with for our analytic purposes, and it is more appropriate to depict certain applications. For example, in the tax inspections regime, the tax payers can be thought of as evading taxes only in discrete amounts (this is the case in real life). However, in the general crime control regime, the intensity of criminal activity should be treated as a continuous variable. Therefore, the continuous strategy setting is regarded as the natural extension of the discrete setting that captures the general picture.

We consider the scenario where the inspectees choose their extend of compliance within an infinite bounded set of strategies, $\Lambda = [0, d]$, generating increasingly illegal profits. Here, we identify the inspectees' available strategies with the corresponding illegal profits that they generate to an undetected violator. We retain the initially introduced framework (i.e., the myopic hypothesis, the best response principle, etc.), adjusting our assumptions and our analysis to the continuous local strategy space when needed. Our intention is to extend the findings of Section 2.3.

The population's state space Λ^N is the set of sequences $x = (\ell_1, \dots, \ell_N)$, where $\ell_n \in \Lambda$ is the n-th inspectee's strategy. This can be naturally identified with the set M_N consisting of the normalized sums of N Dirac measures $(\delta_{\ell_1} + \cdots + \delta_{\ell_N})/N$.

Let the set of probability measures on Λ be $M(\Lambda)$. We rewrite the inspectee's payoff function (2.5), playing for illegal profit $\ell \in \Lambda$ against inspector's invested budget b in the form:

$$\Pi(\ell, b) = r + \big(1 - G(b)\big) \cdot \ell - G(b) \cdot f(\ell). \tag{2.23}$$

Let us introduce the notation δ_x for the sum $\delta_{\ell_1} + \cdots + \delta_{\ell_N}$. We rewrite the inspector's payoff function (2.6) in the form:

$$\Pi(\delta_x/N, b, N) = -b + N \cdot G(b) \cdot \langle f, \delta_x/N \rangle \cdot \frac{\kappa}{N} - N \cdot \big(1 - G(b)\big) \cdot \langle \ell, \delta_x/N \rangle \cdot \frac{\kappa}{N}, \tag{2.24}$$

where for the positive scaling constant, without loss of generality, we set $\kappa = 1$.

Recall the argument we introduced in Section 2.3.2 regarding the inspector's subjective evaluation, which leads to expressions (2.6) and (2.24).

It is rigorously proven in Kolokoltsov [2014] that, given that the initial distribution δ_x/N converges to a certain measure $\mu \in M(\Lambda)$ as $N \to \infty$, the group's strategy profile evolution under the inspector's optimum pressure $\hat{b}$ corresponds to the deterministic evolution on $M(\Lambda)$ solving the kinetic equation $\forall A \subseteq \Lambda$:

$$\dot{\mu}_t(A) = \omega \cdot \int_{z \in A} \int_{y \in \Lambda} \big[\Pi(z, \hat{b}) - \Pi(y, \hat{b})\big] \mu_t(dy) \mu_t(dz), \tag{2.25}$$

or equivalently in the weak form:

$$\frac{d}{dt} \langle g(\cdot), \mu_t \rangle = \omega \cdot \int_{\Lambda^2} g(z) \cdot \big[\Pi(z, \hat{b}) - \Pi(y, \hat{b})\big] \mu_t(dy) \mu_t(dz). \tag{2.26}$$

Recall that Assumption 1 ensures that $\hat{b}$ is well defined. Furthermore, notice that notation $\bar{\ell}, \bar{f}, \bar{\xi}$ introduced in Section 2.3 stands here for the expected values $\langle \ell, \mu \rangle, \langle f, \mu \rangle, \langle \xi, \mu \rangle$, $\forall \mu \in M(\Lambda)$, respectively, where $f = f(\ell)$ and $\xi = \xi(\ell) = \ell + f(\ell)$.

Using this inner product notation and substituting (2.23) into (2.25), the kinetic equation can be written in a symbolic form:

$$\dot{\mu}_t(dz) = \omega \cdot \mu_t(dz) \cdot \big[z - G(\hat{b}) \cdot (z + f(z)) + \langle G(\hat{b}) \cdot (f(\cdot) + \cdot) - \cdot, \mu_t \rangle\big]. \tag{2.27}$$

One can think of (2.25) and (2.27) as the continuous local strategy space equivalents of equations (2.9) and (2.12).

Proposition 3. *A (non-negative) probability measure $\mu \in M(\Lambda)$ is a singular point of* (2.25)*, namely, it satisfies:*

$$\int_{z\in A}\int_{y\in A}\big[\Pi(z,\hat{b})-\Pi(y,\hat{b})\big]\mu_t(dy)\mu_t(dz)=0, \tag{2.28}$$

$\forall A\subseteq\Lambda$, if and only if the inspectees' payoff (2.23) *is constant on the support of μ. Since $M(\Lambda)$ is the set of probability laws on Λ, then supp(μ) cannot be an empty set.*

Proof. We use the inner product notation, $\int_{y\in A}\Pi(y,\hat{b}(\mu_t))\mu_t(dy)=\langle\Pi(\cdot,\hat{b}),\mu_t\rangle$, to rewrite (2.28) in the equivalent form:

$$\int_{z\in A}\big(\Pi(z,\hat{b})-\langle\Pi(\cdot,\hat{b}),\mu_t\rangle\big)\mu_t(dz)=0, \tag{2.29}$$

and the result follows, since (2.29) holds when $\Pi(z,\hat{b})=\langle\Pi(\cdot,\hat{b}),\mu_t\rangle$ for any $z\in A$. □

2.4.1 Linear Fine

We consider a linear fine $f(\ell)=\sigma\cdot\ell,\ell\in\Lambda$, and we extend the definitions (2.11), (2.15) and (2.16) to the continuous strategy setting.

Theorem 6. *Let Assumption 1 hold. Every Dirac measure $\delta_z,\forall z\in\Lambda$ is a fixed point of* (2.25)*. Moreover:*

1. *If $G(\hat{b}(d))>\frac{1}{1+\sigma}$, there is additionally a unique hyperplane of fixed points:*

$$\Theta=\Big\{\mu_\theta\in M(\Lambda)\mid\exists!\,\bar{\ell}\in(0,\min[\ell_c,d]):\langle\ell,\mu_\theta\rangle=\bar{\ell}\ \ \&\ \ G(\hat{b}(\bar{\ell}))=\frac{1}{1+\sigma}\Big\},$$

2. *If $G(\hat{b}(d))=\frac{1}{1+\sigma}$ and $d>\ell_c$, there are additionally infinitely many hyperplanes of fixed points:*

$$\Phi=\Big\{\mu_\phi\in M(\Lambda)\mid\forall\,\bar{\ell}\in[\ell_c,d]:\langle\ell,\mu_\phi\rangle=\bar{\ell}\ \ \&\ \ G(\hat{b}(\bar{\ell}))=\frac{1}{1+\sigma}\Big\}.$$

Proof. Every Dirac measure δ_z for arbitrary $z\in\Lambda$ satisfies (2.28), since by definition it is $\langle\Pi(\cdot,\hat{b}),\delta_z\rangle=\Pi(z,\hat{b})$.

Furthermore, Assumption 1 ensures that $\hat{b}:\Lambda\mapsto[0,\hat{b}(d)]$ is a continuous, non-decreasing, surjective function. In particular, $G(\cdot)$ is strictly increasing in $\hat{b}\in[0,B]$, and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\ell}\in[0,\min[\ell_c,d]]$. Therefore:

1. When $G(\hat{b}(d)) > \frac{1}{1+\sigma}$, there is a unique average value $\bar{\ell} \in (0, \min[\ell_c, d])$ satisfying $G(\hat{b}(\bar{\ell})) = \frac{1}{1+\sigma}$. This unique $\bar{\ell}$ is generated by infinitely many probability measures, μ_θ such that $\langle \ell, \mu_\theta \rangle = \bar{\ell}$, forming a hyperplane of points in $M(\Lambda)$ satisfying (2.25).

2. When $G(\hat{b}(d)) = \frac{1}{1+\sigma}$, and $d > \ell_c$, every average value $\bar{\ell} \in [\ell_c, d]$ satisfies $G(\hat{b}(\bar{\ell})) = \frac{1}{1+\sigma}$. Each one of these infinitely many $\bar{\ell}$ is generated by infinitely many probability measures, μ_ϕ such that $\langle \ell, \mu_\phi \rangle = \bar{\ell}$, forming infinitely many hyperplanes of points satisfying (2.25).

□

We refer to the points δ_z as pure strategy fixed points, to emphasize that they correspond to the group's strategy profiles, such that every inspectee plays the same strategy z. Accordingly, we refer to the points μ_θ, μ_ϕ as mixed strategy fixed points.

Theorem 7. *Let Assumption 1 hold. Consider the fixed points given by Theorem 6. For a linear fine:*

1. *the pure strategy fixed points $\delta_z \notin \Theta, \Phi$, $z \in \Lambda : z \neq d$, are unstable;*

2. *the pure strategy fixed point δ_d is asymptotically stable on the topology of the total variation norm, when $\Phi = \Theta = \emptyset$; otherwise, it is unstable;*

3. *the mixed strategy fixed points μ_θ, μ_ϕ are stable.*

Proof. See Section 2.6.

□

2.4.2 Convex/Concave Fine

Let us extend definitions (2.11) and (2.17) to the continuous strategy setting.

For every $x, y \in \Lambda : x < y$, let us also introduce the auxiliary parameter:

$$q^{x,y} = \frac{x - y}{x - y + f(x) - f(y)}. \tag{2.30}$$

Lemma 2. *For a convex fine, $q^{x,y}$ is strictly decreasing in x for constant y (or vice versa), while for a concave fine, $q^{x,y}$ is strictly increasing in x for constant y (or vice versa). In addition, for a convex fine, $q^{x,y}$ is strictly decreasing in x, y for constant $(y - x)$, while for a concave fine, $q^{x,y}$ is strictly increasing in x, y for constant $(y - x)$.*

Theorem 8. *Let Assumption 1 hold. Then, every Dirac measure $\delta_z, \forall z \in \Lambda$, is a fixed point of* (2.25)*. Moreover, every normalized sum of two Dirac measures $\mu_{x,y} = a_x \cdot \delta_x + a_y \cdot \delta_y, \forall x, y \in \Lambda : x < y,\ a_x + a_y = 1$, is a fixed point of* (2.25)*, uniquely defined for each pair of strategies x, y; $\mu_{x,y}$ exist on condition that they satisfy respectively:*

$$G(\hat{b}(\xi(y))) > q^{x,y} > G(\hat{b}(\xi(x))). \tag{2.31}$$

Proof. Every Dirac measure δ_z, for arbitrary $z \in \Lambda$, satisfies (2.28), since by definition, it is $\langle \Pi(\cdot, \hat{b}), \delta_z \rangle = \Pi(z, \hat{b})$.

Consider a probability measure $\mu^* \in M(\Lambda)$ satisfying (2.28), such that $\langle \ell, \mu^* \rangle = \ell^*$, $\langle f, \mu^* \rangle = f^*$, $\langle \xi, \mu^* \rangle = \xi^*$, $\mu^* \neq \delta_z$. Then, from Proposition 3, μ^* should satisfy $\Pi(x, \hat{b}) = \Pi(y, \hat{b})$ for every pair of $x, y \in supp(\mu^*)$, namely the fraction $q^{x,y}$ should be constant $\forall x, y \in supp(\mu^*)$ and equal to $G(\hat{b}(\xi^*))$.

According to Lemma 2, this is possible only when the support of μ^* contains no more than two elements, namely when it is equivalent to the normalised sum of two Dirac measures, such that $\mu^* = \mu_{x,y} = a_x \cdot \delta_x + a_y \cdot \delta_y$, $\forall x, y \in \Lambda : x < y, a_x + a_y = 1$. Such a probability measure satisfies:

$$G(\hat{b}(\xi_{x,y})) = q^{x,y}, \tag{2.32}$$

where $\xi_{x,y} = \langle \xi, a_x \cdot \delta_x + a_y \cdot \delta_y \rangle = a_x \cdot (x + f(x)) + a_y \cdot (y + f(y))$.

In addition, Assumption 1 ensures that $\hat{b} : [0, \xi(d)] \mapsto [0, \hat{b}(\xi(d))]$ is a continuous, surjective, non-decreasing function. Particularly, $G(\cdot)$ is strictly increasing in $\hat{b} \in [0, B]$, and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\xi} \in [0, \min[\xi(c), \xi(d)]]$. Therefore, for any $\mu_{x,y}$ to exist, namely for (2.32) to hold in each instance, the following condition must hold respectively:

$$G(\hat{b}(\xi(y))) > q^{x,y} > G(\hat{b}(\xi(x))). \tag{2.33}$$

□

We refer to the points $\mu_{x,y}$ as double strategy fixed points, since they correspond to the group's strategy profiles such that the inspectees are distributed between two available strategies.

Theorem 9. *Let Assumption 1 hold. Consider the fixed points given by Theorem 8. Then, for a concave fine:*

1. *the pure strategy fixed points $\delta_z, z \neq d$ are unstable;*

2. *the double strategy fixed points $\mu_{x,y} \neq \mu_{0,d}$ are unstable;*

3. *the double strategy fixed point* $\mu_{0,d}$ *is asymptotically stable;*

4. *the pure strategy fixed point* δ_d *is asymptotically stable on the topology of the total variation norm, when* $\mu_{0,d}$ *does not exist; otherwise, it is unstable.*

Proof. See Section 2.6.

□

2.5 Fixed Points and Nash Equilibria

So far, we have deduced and analysed the dynamics governing the deterministic evolution of the multi-player system we have introduced (assuming the myopic hypothesis for an infinitely large population of indistinguishable, interacting agents). Our intention now is to provide a game-theoretic interpretation of the fixed points we have identified. We work in the context of the discrete strategy setting. The extension to the continuous strategy setting is straightforward.

Let us consider the game Ω_N involving a finite number of $N + 1$ players (N inspectees, one inspector). When the inspector chooses to play strategy $b \in B$ and each of the N inspectees chooses to play the same strategy $i \in S$, then the inspector receives the payoff $\Pi_I(b, x, N)$, and each inspectee receives the payoff $\Pi_i(b)$. Note that the inspectees' collective strategy profile can be thought of as the collection of relative occupation frequencies, $x = (x_i)$.

One then defines an ϵ-approximate Nash equilibrium of Ω_N as a profile of strategies $(\hat{b}(x_N), x_N)$, such that:

$$\hat{b}(x_N) = \operatorname{argmax} \Pi_I(b, x_N, N), \tag{2.34}$$

and for any pair of strategies $i, j \in \{0, \ldots, d\}$, the inequality:

$$\Pi_j(\hat{b}(x_N - e_i/N + e_j/N)) \leq \Pi_i(\hat{b}(x_N)), \tag{2.35}$$

holds up to an additive correction term not exceeding ϵ, where e_i denote the standard basis in $\mathbb{R}^d$.

It turns out that the fixed points identified in Section 2.3 for the discrete strategy setting (and by extension in Section 2.4 for the continuous strategy setting) describe approximate Nash equilibria of Ω_N. We state here the relevant result without a proof. A rigorous discussion can be found in Kolokoltsov [2014]. Recall that for a sufficiently large population N (formally valid for $N \to \infty$), we can approximate the relative occupation frequencies x_i with the probabilities ρ_i obeying (2.12).

Proposition 4. *Under suitable continuity assumptions on Π_i and Π_I:*

1. *any limit point of any sequence x_N, such that $(\hat{b}(x_N), x_N)$ is a Nash equilibrium of Ω_N, is a fixed point of the deterministic evolution* (2.12)*;*
2. *for any fixed point x of* (2.12)*, there exists a $1/N-$Nash equilibrium $(\hat{b}(x_N), x_N)$ of Ω_N, such that the difference of any pair of coordinates of x_N, x does not exceed $1/N$ in magnitude.*

The above result provides a game-theoretic interpretation of the fixed points that were identified by Theorems 2, 4, 6 and 8, independent of the myopic hypothesis. Moreover, it naturally raises the question of which equilibria can be chosen by the agents in the long run. The fixed points stability analysis performed in Sections 2.3 and 2.4 aims to investigate this issue. Furthermore, Proposition 4 states, in simple words, that our analysis and our results are also valid for a finite population of inspectees (recall our initial assumption for an infinitely large N), with precision that is inversely proportional to the size of N.

2.6 Proofs

We make use of the *Hartman-Grobman* theorem, stating that *the local phase portrait near a hyperbolic fixed point is topologically equivalent to the phase portrait of the linearisation*, see, e.g., Strogatz [2014], namely that the stability of a hyperbolic fixed point is preserved under the transition from the linear to the non-linear system. For the non-hyperbolic fixed points we resort to Liapunov's method, see, e.g., Jordan and Smith [2007]. Recall that a fixed point is hyperbolic if all the eigenvalues of the linearisation evaluated at this point have non-zero real parts. Such a point is asymptotically stable if and only if all the eigenvalues have strictly negative real part, while it is unstable (either a source or a saddle) when at least one has strictly positive real part.

Proof of Theorem 3

We rewrite (2.12) in the equivalent form, for $i \in S : i \neq j$, and arbitrary $j \in S$:

$$\dot{\rho}_i(t) = \rho_i(t) \cdot \big(1 - (1+\sigma) \cdot G(\hat{b})\big) \cdot \big(\ell_i - \ell_j + \sum_{n \neq j} \rho_n \cdot (\ell_j - \ell_n)\big). \tag{2.36}$$

The linearization of (2.36) around a pure strategy fixed point p_j, can be written in the matrix form:

$$\dot{\mathrm{p}}(t) = A \cdot \mathrm{p}(t) \tag{2.37}$$

where A is a $d \times d$ diagonal matrix, with main diagonal entries, that is, with eigenvalues:

$$\lambda_i|_{\mathrm{p}_j} = \big(1 - (1+\sigma)\cdot G(\hat{b}(\ell_j))\big)\cdot(\ell_i - \ell_j). \tag{2.38}$$

(i) For the pure strategy fixed point p_0 we get:

$$\lambda_i|_{\mathrm{p}_0} = \ell_i, \tag{2.39}$$

that is strictly positive $\forall\, i \in S : i \neq 0$. Then p_0 is a source.

(ii) For the pure strategy fixed point p_d we get:

$$\lambda_i|_{\mathrm{p}_d} = (\ell_i - \ell_d)\cdot\big(1 - (1+\sigma)\cdot G(\hat{b}(\ell_d))\big), \tag{2.40}$$

that is strictly negative $\forall\, i \in S : i \neq d$ when $G(\hat{b}(\ell_d)) < \frac{1}{1+\sigma} \Leftrightarrow \varTheta = \varPhi = \emptyset$. Then p_d is asymptotically stable.
Otherwise, (2.40) is strictly positive, and p_d is a source.

(iii) For the pure strategy fixed points p_j, $\forall\, j \in S : j \neq 0, d$, (2.38) changes sign between $\ell_j < \ell_i$ and $\ell_j > \ell_i$, when $i \in S : i \neq j$. Then p_j are saddles.

(iv) For the non-isolated, non-hyperbolic mixed strategy fixed points $\mathrm{p}_\theta, \mathrm{p}_\phi$ we resort to Liapunov's method. In particular, we consider the real valued Liapunov function $V \in C^1(\Sigma_{d+1})$:

$$V(\mathrm{p}) = \big(1 - (1+\sigma)\cdot G(\hat{b})\big)^2. \tag{2.41}$$

Differentiating with respect to time, we get:

$$\dot{V}(\mathrm{p}) = -(1+\sigma)\cdot\big(1 - (1+\sigma)\cdot G(\hat{b})\big)\cdot\frac{\partial G}{\partial \hat{b}}\cdot\frac{\partial \hat{b}}{\partial \bar{\ell}}\cdot\frac{\partial \bar{\ell}}{\partial t}. \tag{2.42}$$

From Assumption 1, $G(\cdot)$ is strictly increasing in $\hat{b} \in [0, B]$, and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\ell} \in [0, min[\ell_c, \ell_d]]$. Additionally, differentiating (2.36) with respect to time we get:

$$\frac{\mathrm{d}\bar{\ell}}{\mathrm{d}t} = \big(1 - (1+\sigma)\cdot G(\hat{b}(\bar{\ell}))\big)\cdot\big(\langle \ell, p\rangle^2 - \langle \ell^2, p\rangle\big). \tag{2.43}$$

Overall, we have that $V(\mathrm{p}_\theta) = 0$, $V(\mathrm{p}) > 0$ if $\mathrm{p} \neq \mathrm{p}_\theta$, and $\dot{V}(\mathrm{p}) \leq 0$ for all $\mathrm{p} \in \Sigma_{d+1}$ (respectively for p_ϕ and $\mathrm{p} \neq \mathrm{p}_\phi$). Therefore, according to Liapunov's Theorem, $\mathrm{p}_\theta, \mathrm{p}_\phi$ are stable.

Proof of Theorem 5

We rewrite (2.12) in the equivalent form, $i \in S : i \neq d$:

$$\dot{\rho}_i(t) = g_i(\mathrm{p}) = \big((\ell_i - \ell_d + f_i - f_d) \cdot (q^{i,d} - G(\hat{b})) - \sum_{j \neq d}(\ell_j - \ell_d + f_j - f_d) \\ \times (q^{j,d} - G(\hat{b})) \cdot \rho_j\big) \cdot \rho_i. \tag{2.44}$$

Around an arbitrary fixed point $\mathrm{p}^* = (\rho_i^*)$, the nonlinear system (2.44) is approximated by:

$$\dot{\rho}_i(t) = \sum_{l \neq d} \frac{\partial g_i(\mathrm{p})}{\partial \rho_\kappa}|_{\mathrm{p}=\mathrm{p}^*} \cdot (\rho_\kappa - \rho_\kappa^*), \tag{2.45}$$

which is a linear system with coefficient matrix $A = (a_{i\kappa})$, $\kappa \in S : \kappa \neq d$, with:

$$a_{i\kappa} := \frac{\partial g_i(\mathrm{p})}{\partial \rho_\kappa}|_{\mathrm{p}=\mathrm{p}^*} = \big(-(\ell_\kappa - \ell_d + f_\kappa - f_d) \cdot (q^{\kappa,d} - G(\hat{b}(\xi^*)) \\ + \big(\ell_d - \ell_i + f_d - f_i + \sum_{j \neq d}(\ell_j - \ell_d + f_j - f_d) \cdot \rho_j^*\big) \cdot \frac{\partial G(\hat{b}(\bar{\xi}))}{\partial \rho_\kappa}|_{\bar{\xi}=\xi^*}\big) \cdot \rho_i^* \\ + \delta_{i\kappa} \cdot \big((\ell_d - \ell_i + f_d - f_i) \cdot (G(\hat{b}(\xi^*)) - q^{i,d}) + \sum_{j \neq d}(\ell_j - \ell_d + f_j - f_d) \\ \times (G(\hat{b}(\xi^*)) - q^{j,d}) \cdot \rho_i^*\big), \tag{2.46}$$

where $\xi^* = \langle \xi, \mathrm{p}^* \rangle$. This is the Jacobian Matrix of (2.44) at an arbitrary fixed point $\mathrm{p}^* = (\rho_i^*)$. We use the characteristic equation:

$$det(A - \lambda \cdot I) = 0,$$

to identify the eigenvalues of matrix A for every fixed point.

Let us introduce the notation $E_{i,j}$ for the elementary matrix corresponding to the row/column operation of swapping rows/columns $i \leftrightarrows j$. The inverse matrix of $E_{i,j}$ is itself, namely it is $E_{i,j}^{-1} = E_{i,j}$.

For a pure strategy fixed point p_l, $l \in S : l \neq 1$, first swapping rows $1 \rightleftarrows l$ of A, and then swapping columns $1 \rightleftarrows l$ of the resulting matrix, we obtain the upper triangular matrix:

$$B = E_{1,l}^{-1} \cdot A \cdot E_{1,l}. \tag{2.47}$$

For $l = 1$, the Jacobian matrix A is already an upper triangular matrix. Matrices A and B are similar, that it, they have the same characteristic polynomial and thus the same eigenvalues. Note that the eigenvalues of an upper triangular matrix are precisely its diagonal elements.

Consequently, the eigenvalues of A at a pure strategy fixed point p_l, $l \in S$, are given by:

$$\lambda_i|_{\mathrm{p}_l} = (\ell_l - \ell_i + f_l - f_i) \cdot (G(\hat{b}(\xi_l)) - q^{i,l}) + \delta_{il} \cdot (\ell_i - \ell_d + f_i - f_d) \cdot (G(\hat{b}(\xi_l)) - q^{i,d}). \quad (2.48)$$

For a double strategy fixed point $\mathrm{p}_{m,n}$, $m, n \in S : m < n, m \neq 1, 2, n \neq 2$, swapping rows $1 \rightleftarrows m$ of A, and then swapping in order rows $2 \rightleftarrows n$, columns $1 \rightleftarrows m$, columns $2 \rightleftarrows n$, we obtain the matrix:

$$C = (E_{1,m} \cdot E_{2,n})^{-1} \cdot A \cdot E_{1,m} \cdot E_{2,n}, \quad (2.49)$$

where we have used the inverse matrix product identity.

For $(m, n) = (1, 2)$ the Jacobian matrix A has already the form of C. For $m = 1, n \neq 2$ we need to swap only the n row, n column. Respectively for $m = 2$. Matrices A and C are similar. The characteristic polynomial of C, and thus of A, is:

$$\begin{aligned}
&(a_{mm} - \lambda) \cdot det\big((c_{i\kappa})_{i,\kappa\neq 1} - \lambda \cdot I\big) + (a_{mn} - \lambda) \cdot det\big((c_{i\kappa})_{i\neq 1,\kappa\neq 2} - \lambda \cdot I\big) \\
&= (a_{mm} - \lambda) \cdot \prod_{i\neq m} (a_{ii} - \lambda) + (a_{mn} - \lambda) \cdot (a_{nm} - \lambda) \cdot \prod_{i\neq m,n} (a_{ii} - \lambda) \\
&= \big((a_{mm} - \lambda) \cdot (a_{nn} - \lambda) + (a_{mn} - \lambda) \cdot (a_{nm} - \lambda)\big) \cdot \prod_{i\neq m,n} (a_{ii} - \lambda) \\
&= 0,
\end{aligned} \quad (2.50)$$

where $(c_{i\kappa})_{i,\kappa\neq 1}$ and $(c_{i\kappa})_{i\neq 1,\kappa\neq 2}$ are upper triangular matrices.

Thus, the eigenvalues of A at a double strategy fixed point $\mathrm{p}_{m,n}, m, n \in S : m < n$, are given by:

$$\begin{aligned}
\lambda_i|_{\mathrm{p}_{m,n}} = (\ell_m - \ell_i + f_m - f_i) \cdot (G(\hat{b}(\xi_{m,n})) - q^{i,m}) + (\delta_{in} \cdot \rho_n + \delta_{im} \cdot \rho_m) \\
\times \big((\ell_i - \ell_d + f_i - f_d) \cdot (G(\hat{b}(\xi_{m,n})) - q^{i,d}) \\
+ (\rho_n \cdot (\ell_n - \ell_m + f_n - f_m) + \ell_m - \ell_i + f_m - f_i) \cdot \frac{\partial G(\hat{b}(\bar{\xi}))}{\partial \rho_i}|_{\bar{\xi}=\xi_{m,n}}\big).
\end{aligned} \quad (2.51)$$

Concave Fine;

(i) For the pure strategy fixed point p_0, we get form (2.48):

$$\lambda_i|_{\mathrm{p}_0} = \ell_i + \delta_{i0} \cdot (\ell_d - \ell_i), \quad (2.52)$$

that is strictly positive $\forall\, i \in S : i \neq d$. Then p_0 is a source.

(ii) For the pure strategy fixed points p_l, $l \neq 0, d$, say it is $G(\hat{b}(\xi_l)) > q^{i,l}$ (or $G(\hat{b}(\xi_l)) < q^{i,l}$), $\forall\, i \in S : i \neq d$; then (2.48) changes sign between $i < l$ and $i > l$. Alternatively, say there is some w such that $G(\hat{b}(\xi_l)) > q^{w,l}$, $G(\hat{b}(\xi_l)) < q^{w+1,l}$; then (2.48) indicatively changes sign between $i < w < l$ and $w < i < l$ (or between $l < i < w$ and $l < w < i$). Then p_l, $l \neq 0, d$, are saddles.

(iii) For the double strategy fixed points $\mathrm{p}_{m,n} \neq \mathrm{p}_{0,d}$, (2.51) changes sign, indicatively between $i > n > m$ and $n > i > m$ (since $q^{i,m}$ is strictly increasing in i). Then $\mathrm{p}_{m,n} \neq \mathrm{p}_{0,d}$ are saddles.

(iv) For the double strategy fixed point $\mathrm{p}_{0,d}$, we get from (2.51):

$$\begin{aligned} \lambda_i|_{\mathrm{p}_{0,d}} &= -(\ell_i + f_i) \cdot (G(\hat{b}(\xi_{0,d})) - q^{0,i}) + \delta_{i0} \cdot \rho_0 \cdot \big((\ell_i - \ell_d + f_i - f_d) \\ &\times (G(\hat{b}(\xi_{0,d})) - q^{i,d}) + (\rho_d \cdot (\ell_d + f_d) - \ell_i - f_i) \cdot \frac{\partial G(\hat{b}(\bar{\xi}))}{\partial \rho_i}|_{\bar{\xi}=\xi_{0,d}}\big), \end{aligned} \tag{2.53}$$

that is strictly negative $\forall\, i \in S : i \neq d$ (since $G(\hat{b}(\xi_{0,d})) = q^{0,d}$, $q^{0,i}$ is strictly increasing in i). Then $\mathrm{p}_{0,d}$ is asymptotically stable.

(v) For the pure strategy fixed point p_d, we get from (2.48):

$$\lambda_i|_{\mathrm{p}_d} = (\ell_d - \ell_i + f_d - f_i) \cdot (G(\hat{b}(\xi_d)) - q^{i,d}), \tag{2.54}$$

that is strictly negative $\forall i \in S : i \neq d$ when $\mathrm{p}_{0,d}$ does not exist, namely when $q^{0,d} > G(\hat{b}(\xi_d))$. Then p_d is asymptotically stable.
Otherwise, it is strictly positive $\forall i \in S : i \neq d$, that is, p_d is a source.

Convex Fine;

(i) For the pure strategy fixed point p_0, we get from (2.48):

$$\lambda_i|_{\mathrm{p}_0} = \ell_i + \delta_{i0} \cdot (\ell_d - \ell_i), \tag{2.55}$$

that is strictly positive $\forall i \in S : i \neq d$. Then p_0 is a source.

(ii) For the double strategy fixed points $\mathrm{p}_{m,n}, n \neq m+1$, (2.51) changes sign, indicatively between $i > n > m$ and $n > i > m$ (since $q^{i,m}$ is strictly decreasing in i). Then $\mathrm{p}_{m,n}$ are saddles.

(iii) For the double strategy fixed points $\mathrm{p}_{j,j+1}$, (2.51) is strictly negative $\forall i \in S : i \neq d$ (since $G(\hat{b}(\xi_{j,j+1})) = q^{j,j+1}$, $q^{i,j}$ is strictly decreasing in i). Then $\mathrm{p}_{j,j+1}$ is asymptotically stable.

(iv) For the pure strategy fixed points $\mathrm{p}_l, l \neq 0, d$, (2.48) is strictly negative $\forall i \in S : i \neq d$ when p_l satisfies (2.22), namely when $q^{l-1,l} > G(\hat{b}(\xi_l)) > q^{l,l+1}$. Then p_l is asymptotically stable.
Otherwise (2.48) changes sign (see part (ii) of the proof for a concave fine), that is, p_j is a saddle.

(v) For the pure strategy fixed point p_d, we get from (2.48):

$$\lambda_i|_{\mathrm{p}_d} = (\ell_d - \ell_i + f_d - f_i) \cdot (G(\hat{b}(\xi_d)) - q^{i,d}), \tag{2.56}$$

that is strictly negative $\forall i \in S : i \neq d$ when p_d satisfies (2.22), namely when $q^{d-1,d} > G(\hat{b}(\xi_d))$. Then p_d is asymptotically stable.
Otherwise, it is strictly positive $\forall i \in S : i \neq d$, that is, p_d is a source.

Proof of Theorem 7

(i) From the proof of Theorem (3), we have seen that the pure strategy fixed points $\delta_z \notin \Theta, \Phi$, $z \in \Lambda : z \neq d$, have at least one unstable trajectory.

(ii) For the pure strategy fixed point δ_d, consider the real valued (Liapunov) function $L_1 \in C^1(E)$, where E is an open subset of $M(\Lambda)$, with radius $r < 2$ and center δ_d, such that:

$$L_1(\mu) = d - \langle \ell, \mu \rangle. \tag{2.57}$$

The total variation distance between any two Dirac measures $\delta_x, \delta_y \in M(\Lambda)$ is:

$$d_{TV}(\delta_x - \delta_y) = \sup_{|f| \leq 1} \int f(z)(\delta_x - \delta_y)\mathrm{d}z = 2. \tag{2.58}$$

No Dirac measures are contained in E. Using variational derivatives, we get:

$$\dot{L}_1(\mu) = \big(1 - (1+\sigma) \cdot G(\hat{b}(\bar{\ell}))\big) \cdot \big(\langle \ell, \mu \rangle^2 - \langle \ell^2, \mu \rangle\big). \tag{2.59}$$

When $\Phi = \Theta = \emptyset \Leftrightarrow G(\hat{b}(d)) < \frac{1}{1+\sigma}$, we have that $L_1(\delta_d) = 0$, $L_1(\mu) > 0$ if $\mu \neq \delta_d$, and $\dot{L}_1(\mu) < 0$ for all $\mu \in E \backslash \delta_d$. Then, according to Liapunov's Theorem, δ_d is asymptotically stable.

(iii) For the mixed strategy fixed points μ_θ, μ_ϕ, take the real valued (Liapunov) function $L_2 \in C^1(M(\Lambda))$

$$L_2(\mu) = \big(1 - (1+\sigma) \cdot G(\hat{b}(\bar{\ell}))\big)^2. \tag{2.60}$$

Using variational derivatives, we get:

$$\dot{L}_2(\mu) = 2 \cdot (1+\sigma) \cdot L_2(\mu) \cdot \frac{\mathrm{d}G}{\mathrm{d}\hat{b}} \cdot \frac{\mathrm{d}\hat{b}}{\mathrm{d}\bar{\ell}} \cdot \big(\langle \ell, \mu \rangle^2 - \langle \ell^2, \mu \rangle\big). \tag{2.61}$$

From Assumption 1, $G(\cdot)$ is strictly increasing in $\hat{b} \in [0, B]$, and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\ell} \in [0, \min[\ell_c, d]]$. Hence, we have that $L_2(\mu_\theta) = 0$, $L_2(\mu) > 0$ if $\mu \neq \mu_\theta$, and $\dot{L}_2(\mu) \leq 0$ for all $\mu \in M(\Lambda)$ (respectively for μ_ϕ). Then, according to Liapunov's theorem, μ_θ, μ_ϕ are stable.

Proof of Theorem 9

(i)-(ii) From the proof of Theorem (5), we have seen that the pure strategy fixed points $\delta_z, z \neq d$, and the double strategy fixed points $\mu_{x,y} \neq \mu_{0,d}$, have at least one unstable trajectory.

(iii) For the mixed strategy fixed point $\mu_{0,d}$, consider the real valued (Liapunov) function $U_1 \in C^1(E)$ where E is an open subset of $M(\Lambda)$, with radius $r < 2$ and center $\mu_{0,d}$:

$$U_1(\mu) = \big(q^{0,d} - G(\hat{b}(\bar{\xi}))\big)^2. \tag{2.62}$$

Using variational derivatives we get:

$$\begin{aligned}\dot{U}_1(\mu) = -2 \cdot \frac{\mathrm{d}G}{\mathrm{d}\hat{b}} \cdot \frac{\mathrm{d}\hat{b}}{\mathrm{d}\bar{\xi}} \cdot \big(q^{0,d} - G(\hat{b}(\bar{\xi}))\big) \cdot \Big(\langle \ell^2, \mu \rangle - \langle \ell, \mu \rangle^2 + \langle \ell \cdot f(\ell), \mu \rangle \\ - \langle \ell, \mu \rangle \cdot \langle f(\ell), \mu \rangle - G(\hat{b}(\bar{\xi})) \cdot \big(\langle \ell^2, \mu \rangle - \langle \ell, \mu \rangle^2 + \langle f(\ell)^2, \mu \rangle - \langle f(\ell), \mu \rangle^2 \\ + 2 \cdot (\langle \ell \cdot f(\ell), \mu \rangle - \langle \ell, \mu \rangle \cdot \langle f(\ell), \mu \rangle)\big)\Big).\end{aligned} \tag{2.63}$$

Consider a small deviation from $\mu_{0,d}$:

$$\nu = (1-\epsilon) \cdot (a_0 \cdot \delta_0 + a_d \cdot \delta_d) + \epsilon \cdot \mu, \tag{2.64}$$

where ϵ is small, and $\|\mu\| = 1$. In first order approximation, one can show that:

$$\begin{aligned}\frac{\langle \ell^2, \nu \rangle - \langle \ell, \nu \rangle^2 + \langle \ell \cdot f(\ell), \nu \rangle - \langle \ell, \nu \rangle \cdot \langle f(\ell), \nu \rangle}{\langle \ell^2, \nu \rangle - \langle \ell, \nu \rangle^2 + \langle f(\ell)^2, \nu \rangle - \langle f(\ell), \nu \rangle^2 + 2 \cdot (\langle \ell \cdot f(\ell), \nu \rangle - \langle \ell, \nu \rangle \cdot \langle f(\ell), \nu \rangle)} \\ > (<) q^{0,d} \\ \Rightarrow \langle (f(d) \cdot \ell - d \cdot f(\ell)) \cdot (\ell - a_d + f(\ell) - a_d \cdot f(d)), \mu \rangle > (<) 0,\end{aligned} \tag{2.65}$$

holds, when:

$$\begin{gathered} G\big(\hat{b}\big(a_d \cdot (d+f(d)) + \epsilon \cdot \langle \ell + f(\ell) - a_d \cdot (d+f(d)) \rangle\big)\big) > (<) q^{0,d} \\ \Leftrightarrow \ell + f(\ell) - a_d \cdot (d+f(d)) > (<) 0, \end{gathered} \tag{2.66}$$

where:

$$\langle \xi, \nu \rangle = a_d \cdot (d+f(d)) + \epsilon \cdot \langle \ell + f(\ell) - a_d \cdot (d+f(d)) \rangle. \tag{2.67}$$

For a concave fine, it is:

$$f(d) \cdot \ell - d \cdot f(\ell) < 0, \tag{2.68}$$

and from Assumption 1, $G(\cdot)$ is strictly increasing in $\hat{b}$, and $\hat{b}(\cdot)$ is strictly increasing in $\bar{\ell}$. Then, overall we have that $U_1(\mu_{0,d}) = 0$, $U_1(\mu) > 0$ if $\mu \neq \mu_{0,d}$, and $\dot{U}_1(\nu) < 0$, for any small deviation from $\mu_{0,d}$. Thus, according to Liapunov's theorem $\mu_{0,d}$ is asymptotically stable.

(iv) For the pure strategy fixed point δ_d, consider the real valued (Liapunov) function $U_2 \in C^1(E)$, where E is an open subset of $M(\Lambda)$, with radius $r < 2$ and center δ_d (so that E does not contain any other dirac measures):

$$U_2(\mu) = d - \langle \ell, \mu \rangle. \tag{2.69}$$

Using variational derivatives we get:

$$\begin{gathered} \dot{U}_2(\mu) = -\big(\langle \ell^2, \mu \rangle - \langle \ell, \mu \rangle^2 - G(\hat{b}(\bar{\xi})) \cdot (\langle \ell^2, \mu \rangle - \langle \ell, \mu \rangle^2 \\ + \langle \ell \cdot f(\ell), \mu \rangle - \langle \ell, \mu \rangle \cdot \langle f(\ell), \mu \rangle)\big). \end{gathered} \tag{2.70}$$

When $\mu_{0,d}$ does not exist, namely when $\frac{d}{d+f(d)} > G(\hat{b}(\xi(d)))$, take a small deviation from δ_d:

$$\nu = (1-\epsilon) \cdot \delta_d + \epsilon \cdot \mu, \tag{2.71}$$

where ϵ is small, and $\|\mu\| = 1$. In first order approximation, one can show that:

$$\begin{gathered} \frac{\langle \ell^2, \nu \rangle - \langle \ell, \nu \rangle^2}{\langle \ell^2, \nu \rangle - \langle \ell, \nu \rangle^2 + \langle \ell \cdot f(\ell), \nu \rangle - \langle \ell, \nu \rangle \cdot \langle f(\ell), \nu \rangle} > \frac{d}{d+f(d)} \\ \Rightarrow \langle (\ell - d) \cdot (\ell \cdot f(d) - d \cdot f(\ell)), \mu \rangle > 0, \end{gathered} \tag{2.72}$$

holds, since for a concave fine it is:

$$\ell \cdot f(d) - d \cdot f(\ell) < 0, \tag{2.73}$$

and, for any $\bar{\xi} < \xi(d)$ it is:

$$G(\hat{b}(\bar{\xi})) < G(\hat{b}(\xi(d))). \tag{2.74}$$

Then, overall we have that $U_2(\delta_d) = 0$, $U_2(\mu) > 0$ if $\mu \neq \delta_d$, and $\dot{U}_2(\nu) < 0$ for any small deviation from δ_d within E. Thus, according to Liapunov's theorem, δ_d is asymptotically stable.

Chapter 3

Evolutionary, Mean-Field and Pressure-Resistance Game Modelling of Networks Security

3.1 Introduction

The issue of social security and crime prevention dominantly concerns the modern societies. In the traditional terrain of counter-terrorism, corruption and tax evasion, the corresponding authorities in charge struggle to deal with large populations of increasingly informed violating individuals (this term will be used interchangeably with the terms agents or small players). Reversely, in the recently emerging field of cyber-security, large groups of individuals aim to defend their private computers against a lurking cyber-criminal (bot-net defence). Similar reasoning can be asserted for the citizens of a large city defending against a biological weapon (bio-terrorism). The rapid advance in the means and the speed of interaction, communication and exchange of information has established the individuals' social network as a decisive parameter of their strategic decision making in the above and similar instances. Here we consider agents who are organized in specific social or phenotypic (or even geographical), and behavioural network structures. The central focus of this chapter is to investigate the evolution of the complex process where a (very) large number of interacting individuals, susceptible to engage in or be affected by criminal behaviour, decide their strategies subject to a benevolent, or respectively to a malicious, major player's (this term will be used interchangeably with the term principal) pressure, to their individual optimization criterion, and to their (social) environment's influence.

In the real life scenaria we aim to capture with our approach, it is natural then to distinguish two main dimensions of (network) structure. The first dimension refers to the individuals' objective distribution among different levels of social, bureaucratic, or phenotypic hierarchy, or to their geographical distribution, in general to any finite partition according to their independent characteristics. One can think for example of tax payers of different bands, employees of different grades, or infected computers/individuals of different degrees. The second dimension refers to the agents' distribution among different types of strategy or behaviour, subject (mainly) to the agents' individual control; say for example the level of tax evasion in the field of inspection games, the extent of bribery acceptance in the field of corruption games, or the level of defence against terrorist activity or a malware in the fields of counter terrorism and cyber-security respectively.

Note that our game theoretic approach is developed under the basic idea of a very large number of non-cooperative, interacting agents playing against (i.e under the pressure of) a single major player. In principle, our model belongs to the class of non-linear Markov games, see, e.g., Kolokoltsov [2010], combining under an extended scheme the pressure-resistance, the evolutionary, and the mean-field game approach.

The pressure-resistance terminology was introduced in Kolokoltsov [2014], where ideas captured from evolutionary game theory were extended, including the pressure of a major player on a large group of interacting small players. Here, the pressure-resistance game component refers to the principal's interference that generates transitions solely on the first dimension of structure (e.g. a benevolent director able to promote or downgrade interacting bureaucrats, computers and individuals getting infected or recovering subject to a cyber-criminal's and a bio-terrorist's activity respectively). This approach of major and minor players has also been considered for the analysis of mean-field type models, see, e.g., Bensoussan, Chau and Yam [2016], Carmona and Zhu [2016], Huang [2010].

The evolutionary game component refers to the agents' pairwise interactions, with particular focus on the effect of the established social norms, potentially generating transitions on both dimensions of structure. For a general survey on the literature of population dynamics applications on game theory, that is, on evolutionary game theory, see, e.g., Gintis [2000], Hofbauer and Sigmund [2003], Samuelson [2002], Smith [1988], Szabó and Fath [2007], Taylor et al. [2004], Weibull [1997]. See also Friedman [1991, 1998] for specific application in economics.

The mean-field game (MFG) component refers to the agents' individual optimization controlled by their strategic position on the second dimension of structure, taking into account the entire population's behaviour. This element of 'globally' rational optimization introduces an additional level of complexity compared to Katsikas, Kolokoltsov and Yang [2016], Kolokoltsov [2014], where optimization strictly upon imitation of successful strategies on the basis of binary comparison of payoffs was considered (purely evolutionary approach). MFGs were introduced by Larsy and Lions [2007], by analogy with the mean-field theory in statistical mechanics, and were also introduced independently by Huang, Malham and Caines [2006] as large population stochastic dynamic games. In principle, they represent a natural extension of earlier work in the economics literature under the assumption of infinite number of players, see, e.g. Aumann [1964], Dubey, Mas-Colell and Shubik [1980] for static games, Bergin and Bernhardt [1992], Jovanovic and Rosenthal [1988] for dynamic games. The literature on MFGs is growing fast, see, e.g., Bensoussan, Frehse and Yam [2013], Caines [2013], Cardaliaguet [2010], Carmona and Delarue [2013], Gomes and Saude [2014], Tembine et al. [2009] for a general survey.

Here we shall work in three asymptotic regimes, that is, we shall consider fast execution of the agents' personal decisions, weak binary interactions, and small discounting in time. The need to introduce this ternary asymptotic approach is revealed from the analysis of a similar setting conducted in Kolokoltsov and Bensoussan [2016], where the distribution of infection in a computers network with a malicious software controlled by a cyber-criminal was described by a stationary MFG model with four states. Whilst the three states model describing the distribution of corruption in a population of bureaucrats under the pressure of a benevolent principal that was studied in Kolokoltsov and Malafeyev [2015], is solved explicitly without any asymptotic simplifications, the introduction of a fourth state in Kolokoltsov and Bensoussan [2016] already increases the complexity significantly, such that the need to consider (though not as strongly as we do here) the assumption of large λ (fast decisions execution) is critical to obtain descent solutions.

Similarly, for the even more complex $n \times m$ states model we introduce here, the need to consider the three asymptotic regimes mentioned above becomes obvious. In principle, even without working in these asymptotic regimes one can sometimes obtain explicit but extremely lengthy formulas, not revealing any clearer insights. But also form a practical point of view our asymptotic approach has clear interpretation. Indicatively, an infinitely large transition rate λ implies the natural process of immediate execution of personal decisions as long as they have already been taken, while a vanishingly small discounting δ_{dis} implies a short planning horizon. Both of the models studied in Kolokoltsov and Malafeyev [2015], Kolokoltsov and Bensoussan [2016], and our extended approach here, belong to the category of finite state space mean-field games that were initially considered in Gomes, Mohr and Souza [2010, 2013]. See also Gomes, Velho and Wolfram [2014] for socio-economic applications.

Added to the applications on corruption, and cyber-security, here we introduce the bio-terrorism interpretation, that is, the defence of a population against a biological weapon. The implementation of game theoretic methods to the analysis of terrorism has been vastly developed ever since the 1980s, with game theory allowing the investigation of different instances of strategic interaction (e.g. terrorists vs government, terrorists vs terrorists, terrorists for sponsors, terrorists for supporters), see, e.g., Arce, Daniel and Sandler [2005], Sandler and Arce [2007], Sandler and Siqueira [2009]. The pairing we capture here is civilians vs a bio-terrorist.

We organize Chapter 3 as follows. In Section 3.2 we specify explicitly the time-dependent and stationary MFG consistency problems. In Section 3.3 we solve the stationary problem in our proposed asymptotic regimes, and we show that the identified solution is a stable fixed point of the corresponding evolutionary dynamics. In Section 3.4 we construct the class of time-dependent solutions that stay in a neighbourhood of the identified stationary solution. In the terminology of mathematical economics, this stationary solution represents a turnpike (see, e.g., Kolokoltsov and Yang [2012], Zaslavski [2006]) for the class of time-dependent solutions.

3.2 Formal Model

Let $H = \{1, \cdots, |H| = n\}$ be a finite set characterizing the hierarchical partition of small players inside the environment, say their position in the bureaucratic staircase of an organization. Alternatively, it may describe the extend of individuals' infection to a bio-weapon. Moreover, let $B = \{1, \cdots, |B| = m\}$ be a finite set characterizing the behavioural or strategic partition of agents, say the level of compliance with official regulations, or the degree of protection for PCs/citizens against cyber-criminals/bio-terrorists. Then, the states of an agent are given by ordered pairs of the form (h, b), with $h \in H$, $b \in B$, the finite state space being $S = H \times B$.

Remark 2. *In some cases it is reasonable to include an additional zero state, some kind of a rank-less sink, where no choice of B is available, say a corrupted civil servant suspended from duty without the potential to be bribed, an infected individual put in quarantine, and so forth. Thus, the state space can be either $S = H \times B$ as initially defined above, or $\tilde{S} = H \times B \cup \{0\} = S \cup \{0\}$ as alternatively implied with this comment. We shall stick here to the first instance.*

We distinguish the following three structures. Firstly, the *decision structure* (B, E_D, λ), that is a non-oriented graph with the set of vertices B and the set of edges E_D, where an edge e joins the vertices i and j when an agent is able to switch between states (h, i) and (h, j). Every such transition in B requires certain random λ-exponential time. For simplicity, a single parameter λ is chosen for all possible transitions. As mentioned, we shall mostly look at the asymptotic regime with $\lambda \to \infty$. We take the agents to be homogeneous and indistinguishable, in the sense that their strategies and payoffs depend only on their states, and not on any other individual characteristics. Hence, a decision of an agent in a state (h, b) at any time is given by the *decision matrix* $u = (u_{hb\to h\tilde{b}})$, expressing his intention to switch from b to $\tilde{b}$, for all $\tilde{b} \in B$ such that $\tilde{b} \neq b$. IWe consider agents without mixed strategies, that is, for any state (h, b) the decision vector $(u_{hb\to h\tilde{b}})$ is either identically zero, when the agent does not wish to change strategy, or there exists one strategy $b_1 \neq b$ such that $u_{hb\to hb_1} = 1$, and all the other coordinates of $(u_{hb\to h\tilde{b}})$ being zero, when the agent wishes to change from strategy b to b_1.

Secondly, the *pressure structure* $(H, E_P, q_{jb\to ib})$, that is an oriented graph, where an edge e joins the vertices j and i whenever a major player has the power (or the authority) to upgrade or downgrade the small players from the hierarchy level j to i. In this case, coefficients $q_{jb\to ib}$ represent the rates of such transitions in H, that is, every such transition requires certain $q_{jb\to ib}$-exponential waiting time. In general, these rates may depend on some control of the principal (one can think of some parameter describing the principal's efforts or interference, for example his/her budget). We shall not exploit this version here.

Finally, we consider the *evolution structure* that characterizes the change in the distribution of states due to the agents' pairwise interaction (e.g. through exchange of opinions, fight with competitors, effect of established social norms, and so forth). This can be described by the set of rates $q^s_{s_1\to s_2}$, by which an agent in state s can stimulate the transition of another agent from state s_1 to state s_2. For instance, an honest agent (or even a corrupted one) may help the principal to discover, and therefore punish, the illegal behaviour of a corrupted agent. Note that transitions due to binary interaction can be naturally separated into transitions in B and transitions in H, yielding respectively the *behavioural* and the *hierarchical evolution structures.*

Remark 3. *The scaling $1/N$ for the rates of binary interactions is the standard procedure of making the strength of N^2 (total number of pairs) binary transitions comparable to the strength of N unilateral transitions.*

Here we shall ignore the behavioural element of the evolution structure. That is, we shall assume that transition rates $q^s_{s_1\to s_2}/N$ may not vanish only for two states s_1, s_2 that differ strictly in their h-component. Moreover, since we shall work in the asymptotic regime of small binary interactions, it would be helpful to introduce directly a small parameter δ_{int} discounting the power of these interactions. Then, we shall denote thereafter the corresponding transition rates by $\delta_{int} \cdot q^s_{h_1 b \to h_2 b}/N$.

Remark 4. *The evolutionary transitions in B represent an alternative to the individual transitions described by the decision structure (B, E_D, λ), and can be considered negligible in the limit $\lambda \to \infty$ that we shall look at here. Taking into account a behavioural evolution structure is more appropriate in the absence of a decision structure, which was the case developed in Kolokoltsov and Malafeyev [2015].*

To introduce a more detailed description of our game-theoretic framework, note that the states of the corresponding N players game are the N-tuples of the form, $\{(h_1, b_1), \dots, (h_N, b_N)\}$, where each pair (h_i, b_i) describes each of the N players position on the hierarchy and the behaviour axis respectively. Assuming that each player adopts a decision matrix u, then the system evolves according to the continuous time Markov chain introduced above, with the corresponding transitions rates as were specified. If we further specify the rewards for staying in each state per unit of time, the transition fees/costs for transiting from one state to another, as well as the terminal payoffs corresponding to each state for some finite terminal time, then we shall be working in the setting of a stochastic dynamic game of N players.

As usual in a MFG approach, we are interested in estimating the approximate symmetric Nash equilibria. Assuming indistinguishable agents, the system's state space can be reduced to the set Z_+^{nm} of vectors $n = (n_{ij})$, $i \in H$, $j \in B$, where n_{ij} denotes the number of agents in state (i, j), and $N = \sum_{ij} n_{ij}$ denotes the (constant) total number of agents. Therefore, the initially introduced Markov chain reduces to the Markov chain on Z_+^{nm}, described by the time-dependent generator:

$$\begin{gathered} L_N^t F(n) = \sum_a^n \sum_\beta^m \sum_c^n n_{a\beta} \cdot q_{a\beta \to c\beta} \cdot \big(F(n_{a\beta}^{c\beta}) - F(n)\big) \\ + \sum_a^n \sum_\beta^m \sum_c^n \sum_\gamma^n \sum_k^m n_{a\beta} \cdot \delta_{int} \cdot q^{\gamma k}_{a\beta \to c\beta}/N \cdot n_{\gamma k} \cdot \big(F(n_{a\beta}^{c\beta}) - F(n)\big) \\ + \sum_a^n \sum_\beta^m \sum_c^m n_{a\beta} \cdot \lambda \cdot u_{a\beta \to ac} \cdot \big(F(n_{a\beta}^{ac}) - F(n)\big), \end{gathered} \tag{3.1}$$

where the unchanged values in the arguments of function F on the right-hand side are omitted. Equivalently, in the normalized version the system's state space can be reduced to the subset of the probability simplex $\Sigma^N_{n \times m} \subseteq \mathbb{R}^{n \times m}$, with vectors of the form $x = (x_{ij}) = n/N$, $i \in H$, $j \in B$, where each coordinate will represent now the occupation density (alternatively the occupation probability) of each state (i, j). For the Markov chain on $\Sigma^N_{n\times m}$, generator (3.1) can be rewritten in the equivalent form:

$$\begin{aligned} L^t_N f(x) = \sum_a^n \sum_\beta^m \sum_c^n x_{a\beta} \cdot N \cdot q_{a\beta\to c\beta} \cdot \big(f(x + (e_{c\beta} - e_{a\beta})/N) - f(x)\big) \\ + \sum_{a,c,\gamma}^n \sum_{\beta,k}^m x_{a\beta} \cdot N \cdot \delta_{int} \cdot q^{\gamma k}_{a\beta\to c\beta}/N \cdot x_{\gamma k} \cdot N \cdot \big(f(x + (e_{c\beta} - e_{a\beta})/N) - f(x)\big) \\ + \sum_a^n \sum_\beta^m \sum_c^m x_{a\beta} \cdot N \cdot \lambda \cdot u_{a\beta\to ac} \cdot \big(f(x + (e_{ac} - e_{a\beta})/N) - f(x)\big), \end{aligned} \tag{3.2}$$

where $\{e_{ij}\}$ is the standard orthonormal basis in $\mathbb{R}^{n\times m}$. Assuming, additionally, that f is a continuously differentiable function on $\Sigma^N_{n\times m}$, and taking its Taylor expansion, in the limit of infinitely many agents $N \to \infty$, (3.2) eventually converges to:

$$\begin{aligned} L^t f(x) = \sum_a^n \sum_\beta^m \sum_c^n x_{a\beta} \cdot q_{a\beta\to c\beta} \cdot \big(\frac{\partial f}{\partial x_{c\beta}} - \frac{\partial f}{\partial x_{a\beta}}\big) \\ + \sum_a^n \sum_\beta^m \sum_c^n \sum_\gamma^n \sum_k^m x_{a\beta} \cdot \delta_{int} \cdot q^{\gamma k}_{a\beta\to c\beta} \cdot x_{\gamma k} \cdot \big(\frac{\partial f}{\partial x_{c\beta}} - \frac{\partial f}{\partial x_{a\beta}}\big) \\ + \sum_a^n \sum_\beta^m \sum_c^m x_{a\beta} \cdot \lambda \cdot u_{a\beta\to ac} \cdot \big(\frac{\partial f}{\partial x_{ac}} - \frac{\partial f}{\partial x_{a\beta}}\big), \end{aligned} \tag{3.3}$$

or equivalently to the form:

$$\begin{aligned} L^t f(x) = \sum_{a\neq c}^n \sum_\beta^m \sum_c^n (x_{a\beta} \cdot q_{a\beta\to c\beta} - x_{c\beta} \cdot q_{c\beta\to a\beta}) \cdot \frac{\partial f}{\partial x_{c\beta}} \\ + \sum_{a\neq c}^n \sum_\beta^m \sum_c^n \sum_\gamma^n \sum_s^m (x_{a\beta} \cdot \delta_{int} \cdot q^{\gamma s}_{a\beta\to c\beta} \cdot x_{\gamma s} - x_{c\beta} \cdot \delta_{int} \cdot q^{\gamma s}_{c\beta\to a\beta} \cdot x_{\gamma s}) \cdot \frac{\partial f}{\partial x_{c\beta}} \\ + \sum_\beta^m \sum_c^n \sum_{s\neq\beta}^m (x_{cs} \cdot \lambda \cdot u_{cs\to c\beta} - x_{c\beta} \cdot \lambda \cdot u_{c\beta\to cs}) \cdot \frac{\partial f}{\partial x_{c\beta}}. \end{aligned} \tag{3.4}$$

This is a first order partial differential operator, that generates a deterministic Markov process, whose dynamics are governed by the characteristic equations of L^t:

$$\begin{aligned} \dot{x}_{ij} = \sum_{k\neq j}^m (x_{ik} \cdot \lambda \cdot u_{ik\to ij} - x_{ij} \cdot \lambda \cdot u_{ij\to ik}) + \sum_{a\neq i}^n (x_{aj} \cdot q_{aj\to ij} - x_{ij} \cdot q_{ij\to aj}) \\ + \sum_k^m \sum_{a\neq i}^n \sum_\gamma^n (x_{aj} \cdot \delta_{int} \cdot q^{\gamma k}_{sj\to ij} \cdot x_{\gamma k} - x_{ij} \cdot \delta_{int} \cdot q^{\gamma k}_{ij\to aj} \cdot x_{\gamma k}). \end{aligned} \tag{3.5}$$

These calculations make the following result plausible:

Proposition 5. *Given the Markovian interaction we introduced above consisting of the decision, the pressure-resistance and the evolution structures, if the elements of the matrix-valued function* $x = (x_{ij})$ *denote the occupation probabilities of states* (i, j)*, and* $(u_{i,k\to j})$ *is the decision matrix that may depend on time, then the evolution of* x *is given by system* (3.5)*.*

Remark 5. *For a rigorous explanation (not just the formal description we provide here) of the Markov chain's convergence to the deterministic process given by* (3.5)*, see, e.g., Kolokoltsov [2012].*

The above general structure is rather complicated. To deal effectively with this complexity, one can distinguish two natural simplifying frameworks: (i) the set of edges is ordered and only the transitions between neighbours are allowed, (ii) the corresponding graph is complete, so that all transitions are allowed and have comparable rates. We shall choose the second alternative for B, and the first alternative for H thinking of it as an hierarchy of agents. Moreover, we shall assume that the binary interaction occurs only within a common level in H, ignoring the binary interaction between the agents in different levels of the hierarchy structure. Therefore, for the transition rates $q_{ij\to i+1,j}$ of the pressure structure increasing in $i \in H$, we introduce the shorter notation q^+_{ij}, and for the transition rates $q_{ij\to i-1,j}$ decreasing in $i \in H$, we introduce the notation q^-_{ij}. Accordingly, for the transition rates $\delta_{int} \cdot q^{ik}_{ij\to i+1,j}$ of the hierarchical evolution structure increasing in $i \in H$, we introduce the shorter notation $\delta_{int} \cdot q^{+k}_{ij}$, and for the transition rates $\delta_{int} \cdot q^{ik}_{i\to i-1,j}$ decreasing in $i \in H$, we shall use the notation $\delta_{int} \cdot q^{-k}_{ij}$.

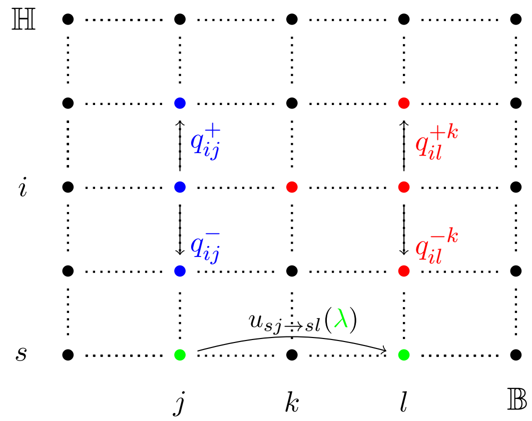


Figure 3.1: The simplified version of our network: only the transitions between neighbours are allowed in H, all transitions are allowed in B, binary interaction occurs only within a common level in H.

Applying the above (geometric) simplifications, the kinetic equations (3.5) reduce to the following system:

$$\begin{aligned}\dot{x}_{ij} = \lambda \cdot \sum_{k\neq j}(u_{ik\to ij}x_{ik} - u_{ij\to ik}x_{ij}) + q^-_{i+1,j}\cdot x_{i+1,j} + q^+_{i-1,j}\cdot x_{i-1,j} - (q^+_{ij} + q^-_{ij})\cdot x_{ij}\\ +\delta_{int}\cdot\sum_{k\in B}(q^{+k}_{i-1,j}\cdot x_{i-1,k}\cdot x_{i-1,j} + q^{-k}_{i+1,j}\cdot x_{i+1,k}\cdot x_{i+1,j} - (q^{+k}_{ij}+q^{-k}_{ij})\cdot x_{ik}\cdot x_{ij}).\end{aligned} \tag{3.6}$$

Note that (3.6) hold only for the internal states (i,j), $i \in H$, $j \in B$, such that $i \neq 1, |H|$, while for the boundary states (i,j) the terms involving downgrading to $i-1$ and upgrading to $i+1$ respectively are omitted. In particular, we consider:

$$q^{+k}_{nj} = q^{-k}_{1j} = 0 \quad , \quad q^+_{nj} = q^-_{1j} = 0. \tag{3.7}$$

Additionally, to simplify further the final explicit calculations, for all $i \in H$, $j \in B$, we shall consider the constraint:

$$q^+_{ij} = q^-_{i+1,j}, \tag{3.8}$$

which can be interpreted as a *detailed balance condition*; it actually asserts that the number of downgrades is compensated in average by the number of upgrades.

Remark 6. *An alternative simple (and analogously manageable) model allows the principal either to move an agent one-step upward in the hierarchy with rates q^+_{ij} and q^{+k}_{ij} respectively, or send an agent directly down to the lowest state with rates q^d_{ij} and q^{dk}_{ij} respectively. In this case the system describing the evolution of occupation densities becomes, for $i \neq 1$:*

$$\begin{aligned}\dot{x}_{ij} = \lambda\cdot\sum_{k\neq j}(u_{i,k\to j}\cdot x_{ik} - u_{i,j\to k}\cdot x_{ij}) + q^+_{i-1,j}\cdot x_{i-1,j} - (q^+_{ij}+q^d_{ij})\cdot x_{ij}\\ +\delta_{int}\cdot\sum_{k\in B}(q^{+k}_{i-1,j}\cdot x_{i-1,k}\cdot x_{i-1,j} - q^{+k}_{ij}\cdot x_{ik}\cdot x_{ij}) - \delta_{int}\cdot\sum_{k\in B} q^{dk}_{ij}\cdot x_{ik}\cdot x_{ij},\end{aligned} \tag{3.9}$$

with an obvious modification for $i = 1$.

To identify the agents' optimal decision vector, we need first to define certain game characteristics such as the state rewards and the transition costs. In particular, we assign the reward w_{ij} per unit of time to an agent for staying in state (i,j), the fee/cost f^B_{kj} for an agent's elective transition from state (h,k) to state (h,j) (which we assume independent of h for brevity), and the fine/cost f^H_j for an agent's enforced transition from state (j, b) to state $(j-1, b)$ (which we assume independent of b for brevity). Let, additionally, $g_{ij} = g_{ij}(t)$ be the payoff corresponding to state (i, j) in the process starting at time t and terminating at time T. Then, for an infinitesimally small time step τ, and assuming that $g(t)$ is continuously differentiable in time, an agent at state (i, j) decides his/her strategy targeting to optimize the expression:

$$\begin{aligned} g_{ij}(t) = \max_u \big\{ \tau \cdot w_{ij} + \tau \cdot \big(\lambda \cdot u_{ij \to ik} \cdot (g_{ik}(t+\tau) - f^B_{jk}) + q^+_{ij} \cdot g_{i+1,j}(t+\tau) \\ + q^-_{ij}(g_{i-1,j}(t+\tau) - f^H_i) + \sum_k^m x_{ik} \delta_{int} (q^{+k}_{ij} g_{i+1,j}(t+\tau) + q^{-k}_{ij}(g_{i-1,j}(t+\tau) - f^H_i))\big) \\ + \big(1 - \tau \cdot (\lambda \cdot u_{ij \to ik} + q^+_{ij} + q^-_{ij} + \sum_k^m x_{ik} \cdot \delta_{int} \cdot (q^{+k}_{ij} + q^{-k}_{ij}))\big) \cdot g_{ij}(t+\tau) \big\}. \end{aligned} \tag{3.10}$$

Remark 7. *Depending on the application we investigate in each instance, the agent's optimum is either to maximize his/her payoff (fitness), or to minimize his/her cost. Here we stick to the first case, thinking of the agents as bribed bureaucrats or defending civilians.*

Taking the Taylor expansion specifically of the term $g_{ij}(t+\tau)$, and omitting terms of order $O(\tau^2)$, the above optimization equation turns into the following form:

$$\begin{aligned} w_{ij} + \frac{\partial g_{ij}(t)}{\partial t} + \max_u \{ \lambda \cdot u_{ij \to ik} \cdot (g_{ik}(t+\tau) - f^B_{jk} - g_{ij}(t)) \} \\ + \sum_k^m x_{ik} \cdot \delta_{int} \cdot (q^{+k}_{ij} \cdot (g_{i+1,j}(t+\tau) - g_{ij}(t)) + q^{-k}_{ij} \cdot (g_{i-1,j}(t+\tau) - f^H_i - g_{ij}(t))) \\ + q^+_{ij} \cdot (g_{i+1,j}(t+\tau) - g_{ij}(t)) + q^-_{ij} \cdot (g_{i-1,j}(t+\tau) - f^H_i - g_{ij}(t)) = 0. \end{aligned} \tag{3.11}$$

In the limit of infinitesimally small time step $\tau \to 0$, (3.11) implies the evolutionary Hamilton-Jacoby-Bellman (HJB) equation, satisfied by the agents' individual optimal payoffs g_{ij}. A rigorous derivation of the HJB equation can be found in every standard textbook on dynamic programming and optimal control, see, e.g., Kamien and Schwartz [1991]. For stochastic dynamic programming, see, e.g., Ross [2014].

The above yields the following result:

Proposition 6. *Given the Markovian interaction we introduced above consisting of the decision, the pressure-resistance and the evolution structure, if $g_{ij} = g_{ij}(t)$ denotes the payoff to an agent at state (ij) in the process starting at time t and terminating at time T, and subject to a given evolution of the occupation density* *vector x given by* (3.6)*, these individual optimal payoffs will satisfy the following evolutionary HJB equation:*

$$\begin{aligned}
&\dot{g}_{ij} + \lambda \cdot \max_u \{u_{ij\to ik} \cdot (g_{ik} - g_{ij} - f^B_{jk})\} + q^+_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^-_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i) \\
&\quad + \delta_{int} \cdot (\sum_{k\in B} q^{+k}_{ij} \cdot x_{ik} \cdot (g_{i+1,j} - g_{ij}) + \sum_{k\in B} q^{-k}_{ij} \cdot x_{ik} \cdot (g_{i-1,j} - g_{ij} - f^H_i)) + w_{ij} = 0.
\end{aligned} \tag{3.12}$$

As above, note that (3.12) hold only for the internal states (i, j), $i \in H$, $j \in B$, such that $i \neq 1, |H|$, while for the boundary states (i, j) the terms involving transitions to $i - 1$ or from $i + 1$ respectively are omitted. Indicatively, for $i = 1$ it is:

$$\begin{aligned}
\dot{g}_{1j} = w_{1j} + \lambda \cdot \max_u \{u_{1,j\to k} \cdot (g_{ik} - g_{1j} - f^B_{jk})\} + q^+_{1j} \cdot (g_{2,j} - g_{1,j}) \\
+ \delta_{int} \cdot \sum_{k\in B} q^{+k}_{1j} \cdot x_{ik} \cdot (g_{2j} - g_{1j}).
\end{aligned} \tag{3.13}$$

We shall consider here the optimization problem of estimating the discounted optimal payoff (alternatively one can look for the average payoff in a long time horizon). Hence, assuming the discounting coefficient δ_{dis} for future payoffs, the evolutionary HJB equation for the discounted optimal payoff $e^{-\delta_{dis}\cdot t} \cdot g_{ij}(t)$ of an agent occupying state (i, j), with any finite planning horizon T, can be written as:

$$\begin{aligned}
&\dot{g}_{ij} + \lambda \cdot \max_u \{u_{ij\to ik} \cdot (g_{ik} - g_{ij} - f^B_{jk})\} + q^+_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^-_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i) \\
&\quad + \sum_{k\in B} \delta_{int} \cdot x_{ik} \cdot (q^{+k}_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^{-k}_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i)) + w_{ij} = \delta_{dis} \cdot g_{ij}(t).
\end{aligned} \tag{3.14}$$

The basic mean-field game consistency problem states that, for some interval $[0, T]$, every agent will benefit from applying the same common control, that is, from adopting the same decision vector. In other words, the MFG consistency condition states that one needs to consider the kinetic equations (3.6) (i.e. the forward system), where the collective control is taken into account, and the evolutionary HJB equations (3.14) (i.e. the backward system), where individual controls are taken into account, as a coupled forward-backward system of equations on a given time horizon $[0, T]$, complemented by some initial condition x^0 for the occupation density vector x, and some terminal condition g_T for the optimal payoff g, such that x, g and the common u solve the aforesaid system. Our aim here is first to identify the solution of the stationary consistency problem, and then to investigate the general time-dependent problem, extending (if possible) our findings for the stationary problem.

As mentioned, we shall work in three asymptotic regimes; fast execution of the agents' personal decisions, weak binary interactions, and small payoff discounting in time.

3.3 Stationary Problem

The stationary MFG consistency problem consists of the stationary HJB equation for the discounted optimal payoff of an agent at state (i, j), with a finite time horizon:

$$\begin{aligned} w_{ij} + \lambda \cdot \max_u u_{i,j\to k} \cdot (g_{ik} - g_{ij} - f^B_{jk}) + q^+_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^-_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i) \\ +\delta_{int} \cdot \sum_{k\in B} x_{ik} \cdot (q^{+k}_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^{-k}_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i)) = \delta_{dis} \cdot g_{ij}, \end{aligned} \tag{3.15}$$

where the evolution given by (3.6) is replaced with the analogous fixed point condition:

$$\begin{aligned} \lambda \cdot \sum_{k\neq j} (u_{i,k\to j} \cdot x_{ik} - u_{i,j\to k} \cdot x_{ij}) + q^-_{i+1,j} \cdot x_{i+1,j} + q^+_{i-1,j} \cdot x_{i-1,j} - (q^+_{i,j} + q^-_{i,j}) \cdot x_{ij} \\ +\delta_{int} \cdot \sum_{k\in B} q^{+k}_{i-1,j} \cdot x_{i-1,k} \cdot x_{i-1,j} + q^{-k}_{i+1,j} \cdot x_{i+1,k} \cdot x_{i+1,j} - (q^{+k}_{ij} + q^{-k}_{ij}) \cdot x_{ik} \cdot x_{ij} = 0. \end{aligned} \tag{3.16}$$

By analogy with the time-dependent problem, for the stationary mean-field game consistency problem one needs to consider equations (3.15), (3.16) as a coupled stationary system. In the asymptotic limit of fast execution of individual decisions, $\lambda \to \infty$, the terms in (3.15), (3.16) containing the transition rates λ should obviously vanish (otherwise they would 'explode' to infinity). For a practical interpretation of this observation, one can think that if the execution of personal decisions is significantly fast, then in a stationary state no agent should be interested in switching his/her strategy. In this case (3.15), (3.16) turn respectively into the following form:

$$\begin{aligned} w_{ij} + q^+_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^-_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i) \\ +\delta_{int} \cdot \sum_{k\in B} x_{ik} \cdot (q^{+k}_{ij} \cdot (g_{i+1,j} - g_{ij}) + q^{-k}_{ij} \cdot (g_{i-1,j} - g_{ij} - f^H_i)) = \delta_{dis} \cdot g_{ij}, \end{aligned} \tag{3.17}$$

and,

$$\begin{aligned} q^-_{i+1,j} \cdot x_{i+1,j} + q^+_{i-1,j} \cdot x_{i-1,j} - (q^+_{i,j} + q^-_{i,j}) \cdot x_{ij} \\ +\delta_{int} \cdot \sum_{k\in B} q^{+k}_{i-1,j} \cdot x_{i-1,k} \cdot x_{i-1,j} + q^{-k}_{i+1,j} \cdot x_{i+1,k} \cdot x_{i+1,j} - (q^{+k}_{ij} + q^{-k}_{ij}) \cdot x_{ik} \cdot x_{ij} = 0, \end{aligned} \tag{3.18}$$

supplemented by the consistency condition:

$$g_{ik} - g_{ij} - f^B_{jk} \leq 0, \tag{3.19}$$

for all $i \in H$, $j, k \in B$, such that $x_{ij} \neq 0$. In fact, the consistency condition (3.19) ensures that all terms in (3.15) and (3.16) including elements of the decision matrix indeed vanish in (3.17) and (3.18) for all the occupied states.

Introducing further the auxiliary notation $\tilde{w}_{ij} = w_{ij} - q^-_{ij} \cdot f^H_i$, (3.17) and (3.18) are written respectively in the form:

$$(-A^T_j + \delta_{dis} - \delta_{int} \cdot E^T_j(x)) \cdot g_{ij} = \tilde{w}_{ij} - \delta_{int} \cdot f^H_i \cdot \sum_{k \in B} q^{-k}_{ij} \cdot x_{ik}, \tag{3.20}$$

and,

$$(A_j + \delta_{int} \cdot E_j(x)) \cdot x_{ij} = 0, \tag{3.21}$$

where the matrices A_j, with the transpose matrix A^T_j, and $E_j(x)$, with the transpose matrix $E^T_j(x)$, are given respectively by:

$$A_j = \begin{pmatrix} -q^+_{1j} & q^-_{2j} & 0 & \dots \\ q^+_{1j} & -q^+_{2j} - q^-_{2j} & q^-_{3j} & \dots \\ \dots & \dots & \dots & \dots \\ \dots & q^+_{n-2,j} & -q^+_{n-1,j} - q^-_{n-1,j} & q^-_{nj} \\ \dots & 0 & q^+_{n-1,j} & -q^-_{nj} \end{pmatrix}, \tag{3.22}$$

and,

$$E_j = \begin{pmatrix} -\sum\limits_k q^{+k}_{1j} x_{1k} & \sum\limits_k q^{-k}_{2j} x_{2k} & 0 & \dots \\ \sum\limits_k q^{+k}_{1j} x_{1k} & -\sum\limits_k (q^{+k}_{2j} + q^{-k}_{2j}) x_{2k} & \sum\limits_k q^{-k}_{3j} x_{3k} & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \sum\limits_k q^{+k}_{n-2,j} x_{n-2,k} & -\sum\limits_k (q^{+k}_{n-1,j} + q^{-k}_{n-1,j}) x_{n-1,k} & \sum\limits_k q^{-k}_{nj} x_{nk} \\ \dots & 0 & \sum\limits_k q^{+k}_{n-1,j} x_{n-1,k} & -\sum\limits_k q^{-k}_{nj} x_{nk} \end{pmatrix}. \tag{3.23}$$

We shall look further for the asymptotic regime with small binary interaction transition rates $\delta_{int} \cdot q^{\pm k}_{ij}$. Therefore, starting with (3.21) we are looking for stationary solutions of the form:

$$x_{ij} = x^0_{ij} + \delta_{int} \cdot x^1_{ij} + O(\delta^2_{int}). \tag{3.24}$$

Substituting (3.24) into (3.21), and equating terms of the same order in δ_{int}^0, δ_{int}^1, we obtain respectively the equations:

$$O(\delta_{int}^0): \quad A_j \cdot x_{ij}^0 = 0, \tag{3.25}$$

$$O(\delta_{int}^1): \quad A_j \cdot x_{ij}^1 + E_j^0 \cdot x_{ij}^0 = 0. \tag{3.26}$$

where the notation E_j^0 corresponds to the matrix E_j containing only elements of order $O(\delta_{int}^0)$ (we use respectively the notation E_j^{0T} for the transpose matrix).

Assumption 2. *Let the detailed balance condition* (3.8) *hold with all transition rates q_{ij}^+ (or all q_{ij}^- respectively) being strictly positive. We shall use then the shorter notation, for $i \in H : i \neq n$, $j \in B$:*

$$q_{ij} = q_{ij}^+ = q_{i+1,j}^-. \tag{3.27}$$

In the linear approximation of vanishing δ_{int}, we end up with an uncoupled system. Since different elements of B are also uncoupled, then, equations (3.25) and (3.26) can be solved separately for any $j \in B$. Looking at the zero order of small evolution transition rates, by (3.25), we have the following result:

Proposition 7. *Let Assumption 2 hold. Then, the rank of A_j is exactly $n-1$, while the kernel of A_j is generated by the following vector:*

$$\begin{gathered} x_{2j}^0 = \frac{q_{1j}^+}{q_{2j}^-} \cdot x_{1j}^0, \quad x_{3j}^0 = \frac{q_{2j}^+}{q_{3j}^-} \cdot \frac{q_{1j}^+}{q_{2j}^-} \cdot x_{1j}^0, \quad \ldots, \quad x_{nj}^0 = \prod_{l=1}^{n-1} \frac{q_{lj}^+}{q_{l+1,j}^-} \cdot x_{1j}^0 \\ x_{1j}^0 = \left(1 + \frac{q_{1j}^+}{q_{2j}^-} + \frac{q_{2j}^+}{q_{3j}^-} \cdot \frac{q_{1j}^+}{q_{2j}^-} + \cdots + \prod_{l=1}^{n-1} \frac{q_{lj}^+}{q_{l+1,j}^-} \right)^{-1} x_j^0. \end{gathered} \tag{3.28}$$

where we have introduced the auxiliary notation $x_j^0 = \sum_i x_{ij}^0$. Specifically, under the detailed balance condition A_j is symmetric, and its kernel generated by (3.28) *is proportional to the uniform distribution, $x_{ij}^0 = x_j^0/n$ for all $i \in H$, $j \in B$, that is, $Ker(A_j)$ is generated by $(1, \cdots, 1)$.*

Proof. Notice that system (3.25) is degenerate, as expected, since we are looking for non-negative solutions satisfying $\sum_j x_{1j}^0 + \cdots + x_{nj}^0 = 1$. Thus, one of the n equations of (3.25) can be discarded, say for example the last one. Rewriting the system of the remaining $(n-1)$ equations by using the first equation, and then adding sequentially to each of the next $(n-2)$ equations their previous one, one eventually obtains the following system:

$$\begin{aligned} q^+_{1j}\cdot x^0_{1j} - q^-_{2j}\cdot x^0_{2j} &= 0\\ q^+_{2j}\cdot x^0_{2j} - q^-_{3j}\cdot x^0_{3j} &= 0\\ &\dots\\ q^+_{n-1,j}\cdot x^0_{n-1,j} - q^-_{nj}\cdot x^0_{nj} &= 0. \end{aligned} \tag{3.29}$$

This has an obvious solution, that is unique up to a multiplier, and is given by (3.28). Alternatively, starting the exclusion from the last equation of (3.29), the solution to (3.25) is given then by:

$$\begin{gathered} x^0_{n-1,j} = \frac{q^-_{nj}}{q^+_{n-1,j}}\cdot x^0_{nj},\;\; x^0_{n-2,j} = \frac{q^-_{n-1,j}}{q^+_{n-2,j}}\cdot\frac{q^-_{nj}}{q^+_{n-1,j}}\cdot x^0_{nj},\;\;\dots,\;\; x^0_{1j} = \prod_{l=1}^{n-1}\frac{q^-_{l+1,j}}{q^+_{l,j}}\cdot x^0_{nj}\\ x^0_{nj} = \left(1+\frac{q^-_{nj}}{q^+_{n-1,j}}+\frac{q^-_{n-1,j}}{q^+_{n-2,j}}\cdot\frac{q^-_{nj}}{q^+_{n-1,j}}+\dots+\prod_{l=1}^{n-1}\frac{q^-_{l+1,j}}{q^+_{l,j}}\right)^{-1}\cdot x^0_j. \end{gathered} \tag{3.30}$$

Given now the detailed balance condition (3.8), and the non-degeneracy established by Assumption 2, one observes from (3.28), or (3.30), that for every strategy $j \in B$ we have:

$$x^0_{1j} = x^0_{2j} = \dots = x^0_{nj} = x^0_j/n.$$

□

We have shown that in the main order of small evolution rates $\delta_{int}\cdot q^{\pm\kappa}_{ij}$, $x^{0*}_{ij} = x^{0*}_j/n$ is a fixed point of the evolution (3.6), along with the common control $u^{com} = (u_{ij\to i\kappa} = 0)$, $\forall i \in H$, $\forall j,\kappa \in B$, that is consistent with condition (3.19), and expresses the instantaneous execution of the agents' personal decisions. Moreover, this will be a stable solution of the stationary system, if $x^{0*}_{ij} = x^{0*}_j/n$ is a stable fixed point of (3.6), for $u^{com} = (u_{ij\to i\kappa} = 0)$, $\forall i \in H$, $\forall j,\kappa \in B$.

Assumption 3. *For technical (computational) purposes only, let the hierarchy and the strategy set be of the same size, i.e.* $|H| = |B| \Rightarrow n = m$.

To conduct a stability analysis, in the asymptotic regimes of large λ and small δ_{int}, let us introduce the auxiliary variables:

$$y_\kappa = x^0_{ij} - x^{0*}_{ij}, \tag{3.31}$$

where $\kappa = i + (j-1) \cdot n$, such that $\kappa \in K = \{1, \dots, n^2 - 1\}$.

Using the above variables, we transform system (3.6) into the non-degenerate linear homogeneous system:

$$\dot{y} = \Lambda \cdot y, \tag{3.32}$$

where Λ is the block matrix:

$$\Lambda = \begin{pmatrix} A_1 & 0 & \dots & \dots & \dots \\ 0 & A_2 & 0 & \dots & \dots \\ \dots & 0 & A_j & 0 & \dots \\ \dots & \dots & 0 & A_{n-1} & 0 \\ \Delta & \dots & \dots & \Delta & D \end{pmatrix}. \tag{3.33}$$

Each matrix Δ has the same non zero entries $-q^-_{nn}$ on its bottom row, while the rest of its elements are equal to zero. Note, as well, that the A_j matrices are of dimension $n \times n$, and each zero matrix to the right of an A_j matrix is of dimension $n \times n \cdot (n-j) - 1$. The $(n-1) \times (n-1)$ matrix D is given by:

$$D = \begin{pmatrix} -q^+_{1n} & q^-_{2n} & 0 & \dots \\ q^+_{1n} & -q^+_{2n} - q^-_{2n} & q^-_{3n} & \dots \\ \dots & \dots & \dots & \dots \\ \dots & q^+_{n-3,n} & -q^+_{n-2,n} - q^-_{n-2,n} & q^-_{n-1,n} \\ \dots & 0 & q^+_{n-2,n} - q^-_{nn} & -q^+_{n-1,n} - q^-_{n-1,n} - q^-_{nn} \end{pmatrix}. \tag{3.34}$$

Applying sequentially (starting with $C_1 \equiv A_1$, setting in the next step $C_1 \equiv A_2$ etc.) the following block matrix formula:

$$\det \begin{pmatrix} C_1 & 0 \\ C_2 & C_3 \end{pmatrix} = \det C_1 \cdot \det C_3,$$

where C_1, C_2, and C_3 are $n \times n$, $m \times n$, and $m \times m$ matrices respectively, the determinant of Λ is given by:

$$\det \Lambda = \det(A_1) \cdot \det(A_2) \cdots \det(A_{n-1}) \cdot \det D. \tag{3.35}$$

We further apply sequentially $n-1$ times the elementary row operation of row addition on every $n \times n$ matrix A_j, starting with row n and adding in each step row i to row $i-1$. Eventually, we transform A_j into a lower triangular matrix of the form:

$$\begin{pmatrix} 0 & 0 & 0 & \dots \\ q^{+}_{1j} & -q^{-}_{2j} & 0 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & q^{+}_{n-2,j} & -q^{-}_{n-1,j} & 0 \\ \dots & 0 & q^{+}_{n-1,j} & -q^{-}_{nj} \end{pmatrix}, \tag{3.36}$$

with a single zero eigenvalue, and $n-1$ negative eigenvalues $-q^{-}_{ij}$, for $i = 2, \dots, n$. Note that since A_j are symmetric matrices (due to the detailed balance condition), the algebraic multiplicity of each of their eigenvalues is equal to the geometric multiplicity.

Regarding the $(n-1) \times (n-1)$ matrix D, and bearing in mind the detailed balance condition, we apply once the elementary row operation of adding row $n-1$ to row $n-2$, and then, we apply sequentially $n-2$ times the elementary column operation of adding column i to column $i+1$, starting with column 1, to eventually transform D into the following lower triangular form:

$$\begin{pmatrix} -q^{+}_{1n} & 0 & 0 & \dots \\ q^{+}_{1n} & -q^{+}_{2n} & 0 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & q^{+}_{n-3,n} & -q^{-}_{n,n} & 0 \\ \dots & 0 & (q^{+}_{n-2,n} - q^{-}_{n,n}) & -q^{-}_{n-1,n} \end{pmatrix}, \tag{3.37}$$

with $n-1$ negative eigenvalues $-q^{+}_{in}$, for $i = 1, \dots, n-1$. In total, we find that matrix Λ has one zero eigenvalue of algebraic multiplicity $n-1$, and $n \cdot (n-1)$ negative eigenvalues. Now it is trivial to transform Λ into a block diagonal matrix, subtracting sequentially from each column i, $\forall i = \{1, \dots, n \cdot n - n\}$, each column j, $\forall j = \{n \cdot n - n + 2, \dots, n \cdot n - 1\}$. For a block diagonal matrix, both the algebraic and the geometric multiplicity of an eigenvalue is given by adding the multiplicities from each block. Then, for the block matrix Λ the algebraic multiplicity of the zero eigenvalue is equal to its geometric multiplicity.

We, thus, have the following result:

Lemma 3. *Let the Assumptions 2, 3 hold. Consider the linear system $\dot{y} = \Lambda \cdot y$ as defined above. The solution to this system, that is the vector $x^{0*}_{ij} = x^{0*}_{j}/n$ given by Proposition 7, is stable (but not asymptotically stable) since Λ has $n \cdot (n-1)$ negative eigenvalues, and a single zero eigenvalue whose algebraic multiplicity equals to its* *geometric multiplicity.*

The third asymptotic regime we shall look at is that of small discounting δ_{dis}. Obviously, no payoff discounting terms appear in the stationary kinetic equations (3.18). Moving to the stationary HJB equation (3.17), or (3.20), initially we are looking for solutions of the form:

$$g_{ij} = g_{ij}^0 + \delta_{dis} \cdot g_{ij}^1. \tag{3.38}$$

Substituting (3.38) into (3.20), and equating terms of zero order in δ_{int} and δ_{dis}, we get the equation:

$$-A_j^T \cdot g_{ij}^0 = \tilde{w}_{ij}, \tag{3.39}$$

In general, equation (3.39) has no (non-degenerate) solution, since (by Proposition 7) the kernel of the symmetric matrix $A_j^T = A_j$ is one dimensional, implying that the image of the transpose matrix A_j^T is $(n-1)$ dimensional (by the rank-nullity theorem). More precisely, equation (3.39) has in general no solution if:

$$(\tilde{w}_{ij}, x_{ij}^0) = \frac{x_j^0}{n} \cdot \sum_i \tilde{w}_{ij} \neq 0. \tag{3.40}$$

Thus, to remain in the non-degenerate regime, we need to introduce additionally the following assumption;

Assumption 4. *For every strategy $j \in B$ the following is true; $\sum_i \tilde{w}_{ij} \neq 0$.*

As a result, we are looking next for solutions of (3.20) in the form of the expansion:

$$g_{ij} = g_{ij}^0 / \delta_{dis} + g_{ij}^1 + g_{ij}^2 \cdot \delta_{dis}. \tag{3.41}$$

Recall that we are looking at the asymptotic regime with small δ_{int} (weak binary interaction), and small δ_{dis} (small payoff discounting). One needs to distinguish clear assumptions on the relationship between the small parameters δ_{int} and δ_{dis}, for a full perturbation analysis. In principle, the following three basic regimes can be naturally identified:

ID$_1$: Interaction is relatively very small, i.e. $\delta_{dis} = \delta$ and $\delta_{int} = \delta^2$.

ID$_2$: Interaction and Discounting are small effects of comparable order, i.e. $\delta_{dis} = \delta_{int} = \delta$.

ID$_3$: Discounting is relatively very small, i.e. $\delta_{int} = \delta$ and $\delta_{dis} = \delta^2$.

We initially concentrate on the ID_1 regime. Substituting (3.41) into (3.20), and equating terms of order δ^{-1}, δ^0, δ^1, we find respectively the following equations:

$$\begin{aligned} A_j^T \cdot g_{ij}^0 &= 0 \\ -A_j^T \cdot g_{ij}^1 + g_{ij}^0 &= \tilde{w}_{ij} \\ -A_j^T \cdot g_{ij}^2 + g_{ij}^1 - E_j^{0T} \cdot g_{ij}^0 &= 0. \end{aligned} \tag{3.42}$$

The first equation in (3.42) tells us that g_{ij}^0 belongs to the kernel of A_j (since $A_j = A_j^T$), that is, for arbitrary constants $a_j \in \mathbb{R}$, we get:

$$g_{ij}^0 = a_j \cdot x_{ij}^0. \tag{3.43}$$

The second equation in (3.42) tells us that $\tilde{w}_{ij} - g_{ij}^0$ belongs to the image of A_j, which coincides with the orthogonal compliment to $Ker(A_j)$, given the identity:

$$Im(A_j) = Ker^{\perp}(A_j^T).$$

Besides, from Proposition 7 we find that the orthogonal compliment to $Ker(A_j)$ is:

$$Ker^{\perp}(A_j) = \{x : \sum_i x_{ij} = 0\}. \tag{3.44}$$

In this case, the fact that $\tilde{w}_{ij} - g_{ij}^0 \in Im(A_j)$ further implies that:

$$\sum_i \tilde{w}_{ij} = \sum_i g_{ij}^0 \Rightarrow \cdots \Rightarrow g_{ij}^0 = \sum_i \tilde{w}_{ij}/n. \tag{3.45}$$

Looking at the third equation in (3.42), and noting that $E_j^{0T} g_{ij}^0 = 0$ for a uniform $g_{\cdot j}^0$, we conclude that $g_{ij}^1 \in Im(A_j)$ as well, that is, $g_{ij}^1 \in Ker^{\perp}(A_j)$. Thus, to identify g_{ij}^1 we need to invert A_j on the reduced $(n-1)$ dimension of $Ker^{\perp}(A_j)$.

Lemma 4. *Let Assumption 2 hold, and let $y \in Ker^{\perp}(A_j)$. Then all solutions z to the matrix equation $A_j \cdot z = y$ are given by the formula:*

$$z_{ij} = z_{1j} - \sum_{a=1}^{i-1} \Big(\sum_{\beta=1}^{a} \frac{y_{\beta j}}{q_{aj}} \Big), \tag{3.46}$$

$\forall i \neq 1$, with arbitrary z_{1j}. There exists a unique solution $z_{\cdot j} \in Ker^{\perp}(A_j)$ specified by:

$$z_{1j} = \sum_{a=1}^{n-1} \Big(\frac{n-a}{n} \cdot \sum_{\beta=1}^{a} \frac{y_{\beta j}}{q_{aj}} \Big). \tag{3.47}$$

Notice that formulae (3.46) and (3.47) yield $z_{ij} = g^1_{ij}$ when $y_{ij} = g^0_{ij} - \tilde{w}_{ij}$. In particular, for g^1_{ij} we find the explicit expression:

$$g^1_{ij} = \sum_{a=1}^{n-1}\Big(\big(\mathbf{1}(i > a)\cdot\frac{n-a-1}{n} + \mathbf{1}(i \le a)\cdot\frac{n-a}{n}\big)\cdot\big(\frac{a}{q_{aj}}\cdot\sum_{\kappa\in H}\frac{\tilde{w}_{\kappa j}}{n} - \sum_{\beta=1}^{a}\frac{\tilde{w}_{\beta j}}{q_{aj}}\big)\Big). \quad (3.48)$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Regarding the consistency condition (3.19), in the main order in small δ it can be written in the equivalent form:

$$\sum_i \tilde{w}_{ik} < \sum_i \tilde{w}_{ij}, \quad (3.49)$$

for all $i \in H$, $k, j \in B$. Given that $\tilde{w}_{\cdot,\cdot}$ does not depend on δ, this leads to the interesting result that in the equilibrium of the asymptotic regime of small δ, only those strategic levels $j \in B$ are occupied (that is, $x^0_j \neq 0$), where the sum $\sum_i \tilde{w}_{ij}$ obtains its maximum. For simplicity, let us further consider the following assumption;

Assumption 5. *There exists a unique behavioural level $b \in B$, such that:*

$$\sum_i \tilde{w}_{ib} > \sum_i \tilde{w}_{ij}. \quad (3.50)$$

Note that Assumption 5 implies that in any equilibrium x^*, with δ sufficiently small, all terms with $j \neq b$ become irrelevant for the analysis.

We, thus, have the following result:

Proposition 8. *Let Assumptions 2, 3, 4 and 5 hold. Consider the ID_1 regime. Then, the solution to the stationary problem described by* (3.17), (3.18) *and* (3.19), *in the main order in small δ, is given by:*

$$x^*_{ib} = x^{0*}_{ib} = 1/n, \quad x^{0*}_{i\kappa} = 0 \;\; \forall \kappa \neq b \in B, i \in H, \quad g_{ib} = \delta^{-1}\cdot g^0_{ib} = \delta^{-1}\cdot\sum_i \tilde{w}_{ib}/n, \quad (3.51)$$

where x^{0*}_{ij} is a stable fixed point of (3.6).

Remark 8. *If we continue in the next order of our perturbation analysis (subsequently in the second next order, and so forth) we can obtain explicit approximate solutions with arbitrary precision.*

Next we consider the ID_2 regime. In this case, we look at the solutions to (3.18) in the next order with respect to small δ. In view of (3.51), we write (3.26) in the form:

$$A_b \cdot x_{ib}^1 + (q_{i-1,b}^{+b} - q_{ib}^{+b} + q_{i+1,b}^{-b} - q_{ib}^{-b})/n^2 = 0, \tag{3.52}$$

where $\sum_i x_{ib}^1 = 0$, and the usual convention for the boundary terms, $i = 1, n$, apply.

Note that the right-hand side of equation (3.52) belongs to $Ker^{\perp}(A_j)$, implying that:

$$\sum_i -(q_{i-1,b}^{+b} - q_{ib}^{+b} + q_{i+1,b}^{-b} - q_{ib}^{-b})/n^2 = 0, \tag{3.53}$$

Moreover, given that $x_{ib}^1 \in Ker^{\perp}(A_j)$, we can identify x_{ib}^1 applying Lemma 4. Formulae (3.46), (3.47) yield $z_{ib} = x_{ib}^1$ when $y_{ib} = -(q_{i-1,b}^{+b} - q_{ib}^{+b} + q_{i+1,b}^{-b} - q_{ib}^{-b})/n^2$.

Regarding the solution to (3.20) in ID_2, substituting (3.41) into (3.20), and equating terms of order δ^{-1}, δ^0, δ^1, we get respectively the following equations:

$$\begin{aligned} A_j^T \cdot g_{ij}^0 &= 0 \\ -A_j^T \cdot g_{ij}^1 + g_{ij}^0 - E_j^{0T} \cdot g_{ij}^0 &= \tilde{w}_{ij} \\ -A_j^T \cdot g_{ij}^2 + g_{ij}^1 - E_j^{0T} \cdot g_{ij}^1 - E_j^{1T} \cdot g_{ij}^0 &= -f_i^H \cdot \sum_k q_{ij}^{-k} \cdot x_{ik}^0, \end{aligned} \tag{3.54}$$

where the notation E_j^1 corresponds to the matrix E_j containing only elements of order $O(\delta_{int})$ (we use respectively the notation E_j^{1T} for the transpose matrix).

The first two equations in (3.54) are identical with the corresponding equations in (3.42) (recall that $E_j^{0T} g_{ij}^0 = 0$ for a uniform g_{ij}^0), and provide the same results expressed through (3.43), (3.45). Looking at the third equation in (3.54), and noting that $E_j^{1T} g_{ij}^0 = 0$, we observe that $(g_{ij}^1 - E_j^{0T} g_{ij}^1 + f_i^H \cdot \sum_k q_{ij}^{-k} \cdot x_{ik}^0) \in Ker^{\perp}(A_j)$, implying that g_{ij}^1 can be uniquely identified through formula (3.46) of Lemma 4, with $z_{ij} = g_{ij}^1$ and $y_{ij} = g_{ij}^0 - \tilde{w}_{ij}$, under the condition:

$$\sum_i (g_{ij}^1 - E_j^{0T} \cdot g_{ij}^1 + f_i^H \sum_k q_{ij}^{-k} \cdot x_{ik}^0) = 0. \tag{3.55}$$

Last we consider the ID_3 regime. Substituting (3.41) into (3.20), but equating now terms of order δ^{-2}, δ^{-1}, δ^0, we get the equations (in analogy to (3.42), (3.54)):

$$\begin{aligned} A_j^T \cdot g_{ij}^0 &= 0 \\ E_j^{0T} \cdot g_{ij}^0 &= 0 \\ -A_j^T \cdot g_{ij}^1 + g_{ij}^0 - E_j^{1T} \cdot g_{ij}^0 &= \tilde{w}_{ij}. \end{aligned} \tag{3.56}$$

Again, the first and the third equations in (3.56) lead to the same results with the first and the second equations in (3.42), namely to (3.43) and (3.45) respectively, while the second equation in (3.56) always holds for a uniform g_{ij}^0.

We, thus, have the following result:

Proposition 9. *The solution to the stationary consistency problem in the main order in small δ in ID_2 and ID_3, is the same with the one identified in Proposition 8 for ID_1.*

3.4 Time-dependent Problem

In principle, the solution to a non-linear Markov game of mean-field type like the one we consider here (on a finite time horizon), defines an ϵ-Nash equilibrium of the corresponding game with a finite number of players, see, e.g., Basna, Hilbert and Kolokoltsov [2014]. Having identified the solution to the stationary MFG consistency problem, we need next to look at the time-dependent consistency problem in order to validate our results for initial/terminal conditions other than those given by the solution of the stationary problem. We further need to investigate the stability of the fixed point x_{ij}^{0*} (see Lemma 3) without assuming that from the very beginning all players apply the same stationary control $u^{com} = (u_{ij\to i\kappa} = 0)$.

For the full time-dependent problem, the HJB equation for the discounted optimal payoff $e^{-\delta_{dis}\cdot t}\cdot g_{ij}(t)$ of an individual at state (i,j) with any planning horizon T is given by (3.14), where now the occupation density vector $x = (x_{ij})$ is also time varying. For definiteness, we shall focus on the ID_1 regime (the same method applies for ID_2, ID_3 regimes). Our aim is to show that by fixing the control $u_{i\alpha\to i\beta} = 0$ in (3.14), $\forall i \in H, \alpha, \beta \in B$, the solution to the occurring system:

$$\begin{aligned}\dot{g}_{i\alpha} + w_{i\alpha} + q_{i\alpha}^{+}\cdot(g_{i+1,\alpha} - g_{i\alpha}) + q_{i\alpha}^{-}\cdot(g_{i-1,\alpha} - g_{i\alpha} - f_i^H)\\ + \sum_{k\in B}\delta_{int}\cdot x_{ik}\cdot(q_{i\alpha}^{+k}\cdot(g_{i+1,\alpha} - g_{i\alpha}) + q_{i\alpha}^{-k}\cdot(g_{i-1,\alpha} - g_{i\alpha} - f_i^H)) = \delta_{dis}\cdot g_{i\alpha}(t),\end{aligned} \tag{3.57}$$

will be consistent, that is, the control $u_{i\alpha\to i\beta} = 0$ will indeed give a maximum in (3.14) in all times.

Fixing the control $u_{i\alpha\to i\beta} = 0$, $\forall i \in H, \alpha, \beta \in B$, is actually equivalent to assuming that:

$$g_{i\beta}(T) - f_{\alpha\beta}^B \le g_{i\alpha}(T). \tag{3.58}$$

Our aim here is to show that starting with a terminal condition belonging to the cone defined by (3.58), we shall stay inside the cone for all $t \le T$. Therefore, it is sufficient to show that on the boundary of this cone the inverted tangent vector of

(3.57) is never directed outside the cone. The necessary condition that needs to be satisfied for this to be true for any boundary point $g_{j\beta} - f^B_{\alpha\beta} = g_{j\alpha}$ is the following:

$$\dot{g}_{j\alpha} - \dot{g}_{j\beta} \leq 0, \tag{3.59}$$

where,

$$\begin{aligned} \dot{g}_{j\alpha} - \dot{g}_{j\beta} &= \delta_{dis} \cdot (g_{j\alpha} - g_{j\beta}) + (w_{j\beta} - w_{j\alpha}) + q^+_{j\beta} \cdot (g_{j+1,\beta} - g_{j\beta}) \\ -q^+_{j\alpha} \cdot (g_{j+1,\alpha} - g_{j\alpha}) &+ q^-_{j\beta} \cdot (g_{j-1,\beta} - g_{j\beta} - f^H_j) - q^-_{j\alpha} \cdot (g_{j-1,\alpha} - g_{j\alpha} - f^H_j) \\ + \sum_{k\in B} \delta_{int} \cdot x_{jk} \cdot \big(& q^{+k}_{j\beta} \cdot (g_{j+1,\beta} - g_{j\beta}) + q^{-k}_{j\beta} \cdot (g_{j-1,\beta} - g_{j\beta} - f^H_j) \\ & -q^{+k}_{j\alpha} \cdot (g_{j+1,\alpha} - g_{j\alpha}) - q^{-k}_{j\alpha} \cdot (g_{j-1,\alpha} - g_{j\alpha} - f^H_j)\big). \end{aligned} \tag{3.60}$$

Substituting g_{ij} from (3.41), and x_{ij} from (3.24) into (3.60) (assuming that f^H_j is independent of δ), and equating terms of similar order, then, in the main order $o(\delta^{-1})$ in small δ, condition (3.59) will be equivalent to (recall that we are in the ID_1 regime):

$$q^+_{j\beta}\cdot(g^0_{j+1,\beta}-g^0_{j\beta})+q^-_{j\beta}\cdot(g^0_{j-1,\beta}-g^0_{j\beta}) \leq q^+_{j\alpha}\cdot(g^0_{j+1,\alpha}-g^0_{j\alpha})+q^-_{j\alpha}\cdot(g^0_{j-1,\alpha}-g^0_{j\alpha}). \tag{3.61}$$

Note that in the main order $o(\delta^{-1})$ in small δ (assuming that $f^B_{\alpha\beta}$ is independent of δ) for the specified boundary point of the cone we get:

$$g^0_{j\beta} = g^0_{j\alpha}, \tag{3.62}$$

while for all the other $i \in H$, such that $i \neq j$, will be:

$$g^0_{i\beta} \leq g^0_{i\alpha}. \tag{3.63}$$

Combining (3.62) and (3.63) we obviously get:

$$g^0_{j\alpha} - g^0_{i\alpha} \leq g^0_{j\beta} - g^0_{i\beta}, \tag{3.64}$$

and rewriting (3.61) in the equivalent form:

$$q^+_{j\alpha}\cdot(g^0_{j\alpha}-g^0_{j+1,\alpha})+q^-_{j\alpha}\cdot(g^0_{j\alpha}-g^0_{j-1,\alpha}) \leq q^+_{j\beta}\cdot(g^0_{j\beta}-g^0_{j+1,\beta})+q^-_{j\beta}\cdot(g^0_{j\beta}-g^0_{j-1,\beta}), \tag{3.65}$$

we check that condition (3.65) is satisfied when $q_{i\alpha} \leq q_{i\beta}$, $\forall i \in H$ (the first term is smaller or equal than the third term, the second term is smaller or equal than the fourth term).

But also for the case when $q_{i\beta} < q_{i\alpha}$, $\forall i \in H$, rewriting (3.64) in the equivalent form:

$$g^0_{i\beta} - g^0_{j\beta} \leq g^0_{i\alpha} - g^0_{j\alpha}, \tag{3.66}$$

we check that (3.61) is satisfied (again the first term is smaller or equal than the third term, the second term is smaller or equal than the fourth term).

We, thus, have the following result:

Proposition 10. *Let Assumptions 2, 3, 4 and 5 hold. Assume additionally,* $\forall \alpha, \beta \in B$*, that:*

$$q_{i\alpha} \leq q_{i\beta} \quad \mathit{or} \quad q_{i\beta} < q_{i\alpha}, \quad \forall i \in H. \tag{3.67}$$

Then, for sufficiently small discounting $\delta_{dis} = \delta$*, and relatively smaller binary interaction coefficient* $\delta_{int} = \delta^2$*, in the main order in small* δ*, for any* $T > t$*, and for any initial occupation probability distribution* $x(t)$*, and any terminal payoffs such that:*

$$g_{i\beta}(T) - f^B_{\alpha\beta} \leq g_{i\alpha}(T),$$

there exists a unique solution to the time-dependent discounted MFG consistency problem such that the control u *is stationary, and is given by* $u_{i\alpha\to i\beta} = 0$, $\forall i \in H$, $\forall \alpha, \beta \in B$*,* $x(s)$ *stays near the fixed point of Proposition 8 as* $s \to T$*, and* $g_{ij}(s)$ *stays near the stationary solution of Proposition 8 (almost for all time), for large* $T - t$*.*

Chapter 4

The Uniformed Patroller

4.1 Introduction

When patrolling a given network against an attack, or infiltration, at an unknown node, there are two plausible scenaria. Either (i) the Patroller may be essentially invisible to the Attacker (e.g. a stealthy drone, a plain-clothes policeman, or a driver of an unmarked car); or (ii) the Patroller may be immediately identifiable by the Attacker (e.g. through his/her uniform - hence our title, or his/her driving a police car). Thus far the literature on patrolling games has concentrated exclusively on the first (invisible Patroller) scenario. To the best of our knowledge, this is the first attempt to model the consequences of the second scenario of a uniformed (noisy) Patroller. Some consequences of this distinction are fairly obvious; for example, a rational mugger will never initiate an attack in a subway car when a uniformed policemen is present. Other consequences, to be explored here, are less obvious.

A game theoretic model of the Attacker-Patroller conflict on a network has recently been the subject of several investigations. Among others, Alpern, Morton and Papadaki [2011] modelled the problem as a zero-sum game between a Patroller and an Attacker, where the Attacker can attack a chosen node in a chosen time period and the Patroller hopes to intercept the attack in time, by following a chosen walk on the network. Their common payoff is the probability that the attack is intercepted. Formally, the Attacker chooses a node i to carry out an attack for a time interval J (a finite sequence of $m \geq 2$ consecutive periods of time); accordingly, the attack is intercepted if the Patroller is at node i at some time period within the interval J. Here, we keep these dynamics and payoff the same. However, to partially simplify the problem we restrict the Patroller to Ergodic Markovian strategies, such that every time the Patroller arrives at a node of the network he/she leaves it by the same distribution over its neighbours (and the node itself).

Note that in the original game-theoretic formulation presented above, the Attacker's strategy (i, J) cannot depend in any way on the Patroller's locations at any time prior to the interval J. However, in many real world situations of this type, the Patroller is identifiable; he/she might be wearing a uniform, driving in a marked car, or making a characteristic noise. Therefore, as we assume in our model presented here, the Attacker can go to the node he/she wishes to attack, and wait there prior to his/her attack. The Attacker can observe when the Patroller is present at the chosen attack node, and when he/she is not. Once the Patroller leaves the attack node, the Attacker only knows that the Patroller is not there; namely, he/she cannot see the Patroller from a distance (vision is limited to the attack node). This assumption defines a game $U(Q, m)$, where Q is the network to be defended, and m is the difficulty of the attack or infiltration as measured in terms of the time required.

The Attacker's pure strategy can now be defined as an ordered pair (i, d), such that he/she goes to node i and initiates his/her attack after the Patroller has been there first and, subsequently, has been away for d consecutive periods (delay time). However, since the Patroller is identifiable (e.g. wearing a uniform), his/her Ergodic Markovian strategy can be assumed to be known to the Attacker; therefore, we adopt a Stackelberg approach where the Patroller is the first to move. This is actually a common assumption in patrolling problems. We solve this game for small values of m and several families of networks: star networks, line networks, circle networks and star-in-circle networks. In particular, both line and circle networks can be interpreted as perimeters of regions to be defended against infiltration. As such, these are models of border defence by uniformed patrollers, thus modifying the border patrol game introduced by Papadaki et al. [2016].

The motivation for the *Uniformed Patroller Problem* originates in the 1970's, when a uniformed policeman was assigned to every subway train (consisting of ten cars, thus being equivalent to the line network L_{10}) in New York City. In the first months at least, the policeman patrolled in a back and forth motion, from car *1* to car *10* and backwards. This patrol was evidently quite foolish as, in our notation, the attack strategy $(1, 1)$ (attacking an end car as soon as the Patroller has left it) would guarantee a win even if the difficulty m of the attack (i.e. the attack duration) was as large as 17. Finding optimal patrolling strategies in this context remains an open issue until now, over thirty years later. Indicatively, note that if the Patroller on this train is following a random walk and the time required for a mugging is for example $m = 2$, it is not optimal to attack as soon as the Patroller leaves your end car ($d = 1$), as he will catch you with probability $1/2$. It is clearly better to wait for some larger number of periods d in which the Patroller is away.

Patrolling related problems have been studied for long, see, e.g., Morse and Kimball [1951], but almost exclusively from the Patroller's point of view. A game theoretic approach, modelling an adversarial Attacker who wants to infiltrate or attack a network at a node of his/her choice, has only recently been introduced by Alpern, Morton and Papadaki [2011]. The techniques developed there were later applied to the class of line networks by Papadaki et al. [2016], with the interpretation of patrolling a border. Other research following a similar reasoning includes Lin et al. [2013] for random attack times, Lin, Atkinson and Glazebrook [2014] for imperfect detection, Hochbaum, Lyu and Ordez [2014] on security routing games, and Basilico, De Nittis and Gatti [2015] for uncertain alarm signals. See also Baykal-Gürsoy et al. [2014] on infrastructure security games. Earlier work on patrolling a channel/border with different paradigms, includes Washburn [1982, 2010], Szechtman et al. [2008], Zoroa N, Fernández-Sáez and Zoroa P [2012], and Collins et al. [2013]. The related problem of ambush is studied by Baston and Kikuta [2004, 2009], while an artificial intelligence approach to patrolling is given by Basilico, Gatti and Amigoni [2012]. Applications to airport security and counter terrorism, are given respectively by Pita et al. [2008], and Fokkink and Lindelauf [2013]. The problem introduced here, of patrolling a network where the Attacker can identify the Patroller only when he/she is in close proximity (e.g. when both being at the same node), appears to be new.

We organize Chapter 4 as follows. In Section 4.2 we explain our model giving a specification of the uniformed Patroller problem on an arbitrary finite network. In Section 4.3 we solve the game for the star network S_n with a central node connected to n ends. It turns out that the Attacker must attack at an end after delaying for $d = 2$ periods; accordingly, the Patroller must decide with what probability to stay at the center when being there. In Section 4.4 we solve the game for the line graph L_n with n nodes, for $n = 4, 5$. Here an attack should be executed at an end node. The optimal delay and the (Ergodic) Markovian patrol depend on the attack difficulty m and the number of nodes. In Section 4.5 we analyse the game on the circle network C_n with n nodes. A point of interest here is that the Patroller can intercept an attack either returning from the direction that he/she left the attack node, or from the opposite direction around the circle. In Section 4.6 we consider a hybrid network consisting of a circle network with a center that is connected to all nodes of the circle. Finally, in Section 4.7 we allow non-Markovian patrols, and identify links between our game and the so-called 'spy games', where both players can see each other, or surveillance problems, where an unmanned aerial vehicle (UAV) must travel across a network being away from each node for no longer than a specified time.

4.2 Formal Model

An attack strategy is a pair (i, d), where i is a node of a given finite network Q, and d is the number of consecutive time periods that the Patroller (e.g. the subway train policeman) must be away from node i for the Attacker to initiate his/her attack. To illustrate the waiting parameter d (we use for this alternately the terms *waiting time* and *delay*), suppose that the presence (at the Attacker's chosen node i) of the Patroller is indicated by a 1, respectively his/her absence is indicated by a 0. Suppose further that the Attacker waits until the Patroller arrives at the attack node, and after that his/her presence-absence sequence is say 11010110 . . . . If the Attacker's waiting time is for example $d = 2$, then his/her decision as to whether or not to attack in each subsequent period is illustrated below:

1	11	110	1101	11010	110101	1101010	11010100
wait	*wait*	*wait*	*wait*	*wait*	*wait*	*wait*	*attack!*

Table 4.1: The Attacker attacks after $d = 2$ periods of the Patroller's absence

If for example the attack difficulty is $m = 3$, then the attack will be successful only if the Patroller's sequence continues with two more 0's, that is, 1101010000 . . . Note that the attack can begin in the same period that the Patroller's absence has been observed. Thus, we take $m > 1$ as otherwise the Attacker could win simply by attacking as soon as the Patroller is not present at his/her node. The above discussion explains the critical parameter d from the Attacker's point of view, who simply observes the alternation of Patroller's presence or absence at his/her chosen node. To illustrate the attack mechanism in the context of both players, take the network Q as the line graph L_4 with four nodes, and consider the dynamics presented in Figure 4.1, where the time axis is drawn horizontally and the line graph is drawn vertically (the nodes are labelled 1 to 4 from top to bottom).

Suppose that the Attacker chooses to attack node *2*, of difficulty $m = 2$, with delay $d = 2$, and say the Patroller is at this node at time $t = 0$. Hence, the Attacker will remain at node *2* indicated by the horizontal red line in Figure 4.1. We consider two potential patrols adopted by the Patroller on the line network L_4; the walk $w_1 = (2, 1, 2, 1, 1, 2, \dots)$ drawn in green on top, and the walk $w_2 = (2, 3, 4, 3, 2, 3, \dots)$ drawn in blue on the bottom. Take first the Attacker's response to w_1. At time $t = 2$ the Attacker resets his/her waiting clock to zero, but at time $t = 4$ since the Patroller is still away after $d = 2$ periods, he/she initiates the attack that lasts for the time interval $\{4, 5\}$ indicated by a thick green horizontal line. However, since the Patroller is back at node *2* at time $t = 5$, the attack is intercepted (the green patrol intersects the thick green horizontal attack). Next, consider the Attacker's response to w_2. At time $t = 2$ he/she begins the attack that lasts for the time interval $\{2, 3\}$ indicated by a thick blue horizontal line. Since the Patroller is not back at node *2* after $m = 2$ periods, the attack is not intercepted (the blue patrol is disjoint from the thick blue horizontal attack). Thus, against the attack strategy (node *2*, $d = 2$), the patrol w_1 wins for the Patroller, while the patrol w_2 loses.

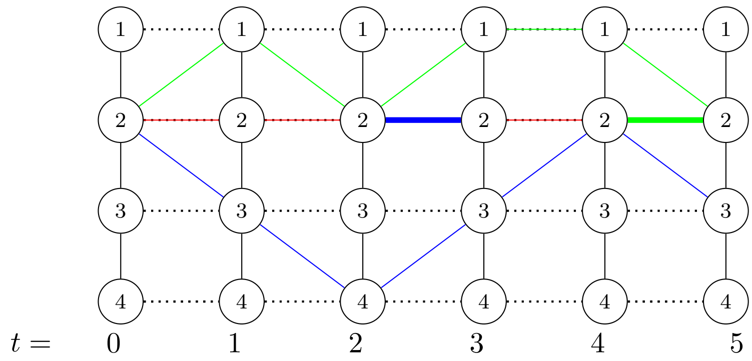


Figure 4.1: Example of Attacker-Patroller dynamics for the line graph L_4.

The reader might as well think that the Attacker has additional strategies which are not specified by our restriction simply to pairs (i, d). For example he/she could initiate the attack after say three consecutive 0's when he/she initially arrives at node i, without necessarily first waiting for the Patroller to visit node i and then counting three consecutive 0's. However, we show that the Attacker can always gain at least as good an outcome using a strategy of the type (i, d). In particular, since we assume that the Patroller has been patrolling for an arbitrarily long time before the period when the Attacker arrives at node i, then we could tell the Attacker, for free, the total number of 0's at node i since the last 1 (including the three 0's the Attacker has witnessed). Say for example this total is 7. If the Attacker chooses to ignore this new information, and attack as planned, his/her expected payoff is the same as that of the strategy $(i, 7)$. On the other hand, if the Attacker changes his/her strategy and decides not to attack, then he/she is consistent with some strategy (i, d).

4.3 The Star Network

The network S_n is the star with a single center connected to n end nodes. We restrict the Patroller to Markovian strategies that reflect from the ends with probability s, from the center c go to each end with equal probability p, and remain at the center with probability $r = 1 - n \cdot p$. This setting simplifies the game by introducing a single parameter family of patrolling strategies. We begin by assuming that the attack takes place at an end node, which we denote by e, and then we show that the Patroller should reflect from the ends ($s = 1$). Note that since we have taken $m \geq 2$, reflecting from the ends will further imply that the Attacker should never attack at the center c because in that case the Patroller will never be away from it for two consecutive periods (i.e. will always intercept such an attack).

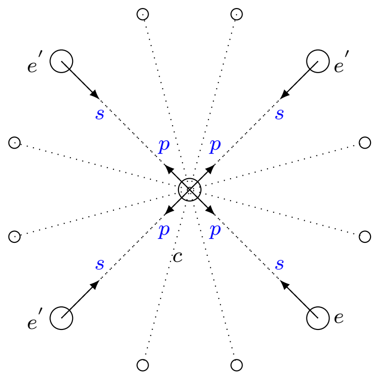


Figure 4.2: The star network S_n with a central node connected to n ends.

4.3.1 Attack duration $m = 2$

Initially we assume that an attack takes $m = 2$ periods. In this case, an attack at node e cannot be intercepted if it starts when the Patroller is at another end node $e^{'} \neq e$, but only if the Patroller is at the center. Therefore, an attack at node e will be intercepted with probability $q \cdot p$, where q is the probability that the Patroller is at the center c at the beginning of the attack, which in turn implies that the Attacker should choose the waiting time d so as to minimize q.

We wish to calculate how the probability q that the Patroller is at c changes over time, as the Patroller continues to be away from the Attacker's chosen node e. Thus, suppose that at some period t the Patroller is not at e, but he/she is either at c with probability q, or at one of the end nodes other than e with probability $1 - q$. Then, in the following period $t + 1$ the Patroller will be either at c with probability $q \cdot r + (1 - q) \cdot s$, or he/she will be at node e with probability $q \cdot p + (1 - q) \cdot 0$.

Hence, conditional on the Patroller not being at node e, the probability that he/she is at the center c is given by:

$$f(q, s) = \frac{q \cdot r + (1 - q) \cdot s}{1 - p \cdot q}. \tag{4.1}$$

Fraction (4.1) is increasing in s, and since it is $s \leq 1$, then we have that $f(q,s)$ is maximized for $s = 1$ at:

$$f(q) = f(q,1) = \frac{q \cdot r + (1-q)}{1 - p \cdot q} = \frac{1 - n \cdot p \cdot q}{1 - p \cdot q}. \tag{4.2}$$

Additionally, since from equation (4.2) we find that:

$$f^{'}(q) = -\frac{p \cdot (n-1)}{(1 - p \cdot q)^2} < 0,$$

then $f(q)$ is decreasing and therefore minimized for $q = 1$, with $f(1) = \hat{q}$, where it is:

$$\hat{q} = \frac{(1 - n \cdot p)}{(1 - p)}. \tag{4.3}$$

The Attacker can obtain this minimum probability $\hat{q}$ of the Patroller being at the central node c at the beginning of the attack, by initiating the attack on the second period that the Patroller is away from his/her chosen node e, that is, by adopting the waiting time $d = 2$. Notice that the optimal Attacker's strategy $(e, 2)$ does not depend on p, namely the initial assumption that the Attacker knows p is not necessary in this case. The attack $(e, 2)$ will be intercepted if the Patroller is at c in its first period and he/she goes to e in its second period, that is, with probability:

$$a_2(n,p) = \hat{q} \cdot p = b(p) = \frac{(1 - n \cdot p) \cdot p}{1 - p}. \tag{4.4}$$

For a given star network S_n, the Patroller will choose the value of p in order to maximize the interception probability (4.4). Figure 4.3 shows the variation of the interception probability $a_2(n,p)$ with p, respectively for $n = 2, \ldots, 8$ arcs in the star network. Recall that $p \in [0, 1/n]$ is the probability with which the Patroller moves to an end node when he/she is at the center c.

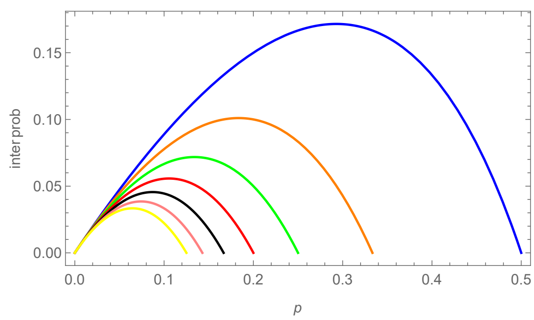


Figure 4.3: The interception probability $a_2(n,p)$, for $n = 2(\text{blue}), \ldots, 8(\text{yellow})$.

Indicatively, for the $n = 2$ arcs star (which is equivalent to a line network with three nodes), the Patroller's optimal walk is to move from the center towards each end with probability about 0.3, and remain at the center with probability about 0.4. For the $n = 3$ arcs star, the Patroller's optimal walk is to move towards each of the three end nodes with probability about 0.2, and remain at the center with probability about 0.4, etc.. We can be more precise about these Patroller's optimal strategies.

In particular, the optimal value $\hat{p} = \hat{p}(n)$ for p depends on n and can be found by solving the first order equation:

$$a_2^{'}(p) = \frac{(n \cdot p^2 - 2 \cdot n \cdot p + 1)}{(1-p)^2}. \tag{4.5}$$

This can be simply written in the following form:

$$n \cdot p^2 - 2 \cdot n \cdot p + 1 = 0, \tag{4.6}$$

giving the optimal values:

$$\hat{p} = 1 - \frac{\sqrt{n \cdot (n-1)}}{n}, \tag{4.7}$$

and

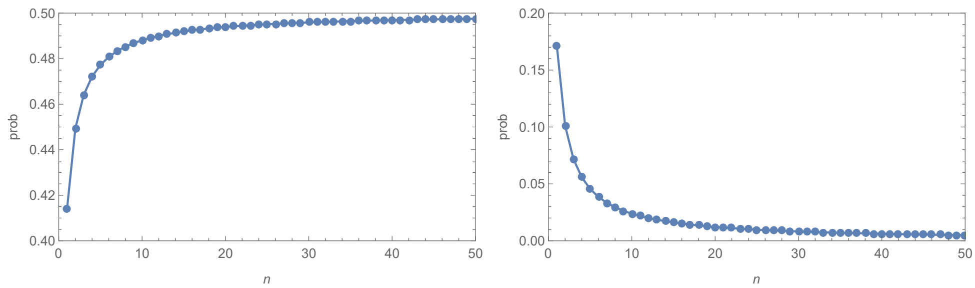
$$\hat{r} = \sqrt{n \cdot (n-1)} - (n-1). \tag{4.8}$$

Equations (4.7), (4.8) show respectively that the optimal probability $\hat{p}$ is asymptotic to $1/(2 \cdot n)$, while the optimal probability $\hat{r}$ goes asymptotically to $1/2$.

The value V of the game is given by

$$V = a(n, \hat{p}) = (2 \cdot n - 1) - 2 \cdot \sqrt{n \cdot (n-1)}. \tag{4.9}$$

In Figure 4.4 we plot the optimal patrol in terms of the probability of remaining at the center, and the corresponding value of the game, for increasing number of nodes.



(a) patroller's optimal walk $\hat{r}(n)$ (b) interception probability $V(n)$

Figure 4.4: Optimal values for $m = 2$.

We can pull together the above results into the following proposition;

Proposition 11. *Consider the Uniformed Patroller Problem on the star network S_n consisting of n arcs, and an attack difficulty of $m = 2$. The optimal Patroller's strategy is to reflect off the end nodes and to remain at the central node with probability:*

$$\hat{r}(n) = \sqrt{n \cdot (n-1)} - (n-1). \tag{4.10}$$

Accordingly, the optimal Attacker's strategy is to locate at a random end node and initiate the attack in the second period that the Patroller is away (i.e. $d = 2$). The interception probability under these strategies, namely the value of the game, is:

$$V(n) = (2 \cdot n - 1) - 2 \cdot \sqrt{n \cdot (n-1)}. \tag{4.11}$$

Remark 9. *We observe that Proposition 11 suggests an attack strategy which does not require any prior knowledge of the Patroller's Ergodic Markovian strategy. That is, the pair of strategies we mention there determines a Nash equilibrium of the uniformed Patroller game. This is indeed a significantly stronger result compared to what we were seeking for when initially considered a Stackelberg approach. Note that this Nash property holds in our later results as well, though we make no claim here for the general existence of a Nash equilibrium in our game for all networks.*

4.3.2 Attack duration $m = 3$

Here we investigate the case when the attack takes $m = 3$ periods. Again, suppose that the attack takes place at the end node e, and at a time period when the conditional probability that the Patroller is at the center provided that he/she is not at node e is given by q. Then, the attack will be intercepted in three different instances corresponding to the Patroller's location at the three periods of the ongoing attack;

$$ce_{-} \quad , \quad cce \quad , \quad e^{'}ce$$
(where $e^{'} \neq e$ is any other end node),

The cumulative probability of these three distinct events, namely the interception probability of the attack, is given by:

$$q \cdot p + q \cdot r \cdot p + (1-q) \cdot s \cdot p = p \cdot \big((1 + r - s) \cdot q + s\big), \tag{4.12}$$

which is increasing in q, since $p < 1/n$. Therefore, as for the case when $m = 2$, the Attacker should begin the attack when the conditional probability q that the

Patroller is at the center given that he/she is not at the attack end node is minimized.

We have seen from the same analysis for $m = 2$ in Section 4.3.1 that this occurs when $d = 2$, where the minimized $q = \hat{q}$ is given by (4.3). Then, substituting equation (4.3) into (4.12) we get:

$$p \cdot \Big((1 + r - s) \cdot \frac{r}{1-p} + s\Big) = \Big(1 - \frac{1-np}{1-p}\Big) \cdot s - \frac{r}{p-1} \cdot (r+1). \tag{4.13}$$

Since the coefficient of s on the right hand of (4.13) is evidently positive, it follows that the Patroller maximizes the interception probability (4.13) by taking $s = 1$ (i.e. reflection), so that the interception probability (4.12) simply becomes:

$$q \cdot p + q \cdot r \cdot p + (1-q) \cdot p. \tag{4.14}$$

Conclusively, substituting the optimal values $q = \hat{q}$ and $s = 1$, the interception probability (4.12) becomes:

$$a_3(n,p) = p \cdot \Big(1 + \frac{(1 - n \cdot p)^2}{1-p}\Big). \tag{4.15}$$

Lemma 5. *For $n > 1$, the interception probabilities $a_3(n,p) = g_n(p)$ are increasing for $0 \leq p \leq \frac{1}{n}$.*

Proof. We have from (4.15) that:

$$g_n^{'}(p) = \frac{h_n(p)}{(1-p)^2}, \tag{4.16}$$

where $h_n(p)$ is the polynomial:

$$h_n(p) = -2 \cdot n^2 \cdot p^3 + 3 \cdot n^2 \cdot p^2 + 2 \cdot n \cdot p^2 - 4 \cdot n \cdot p + p^2 - 2 \cdot p + 2. \tag{4.17}$$

To prove our claim we must show that $h_n(p)$ are positive on $[0, \frac{1}{n}]$. The polynomials $h_n(p)$ are convex on $[0, \frac{1}{n}]$, since:

$$h_n^{''}(p) = 6 \cdot n^2 + 4 \cdot n - 12 \cdot n^2 \cdot p + 2 \geq 6 \cdot n^2 + 4 \cdot n - 12 \cdot n + 2, \tag{4.18}$$

where $p \cdot n \leq 1$. Then, we have for $n \geq 1$ that:

$$h_n^{''}(p) \geq 6 \cdot n^2 - 8 \cdot n + 2 \geq 0. \tag{4.19}$$

Moreover, the polynomials $h_n(p)$ have a unique minimum on $[0, \frac{1}{n}]$ when their first derivative satisfies:

$$h_n^{'}(p) = -6 \cdot n^2 \cdot p^2 + 6 \cdot n^2 \cdot p + 4 \cdot n \cdot p - 4 \cdot n + 2 \cdot p - 2 = 0, \tag{4.20}$$

or equivalently, when probability p is equal to:

$$\bar{p} = \frac{2 \cdot n + 1}{3 \cdot n^2}, \tag{4.21}$$

which is evidently less than $\frac{1}{n}$, since for $n \geq 1$ it is:

$$\frac{1}{n} - \frac{2 \cdot n + 1}{3 \cdot n^2} = \frac{n - 1}{3 \cdot n^2}. \tag{4.22}$$

Conclusively, we find for $n \geq 1$ that:

$$h_n(\bar{p}) = \frac{(n-1)^2 \cdot (18 \cdot n^2 + 8 \cdot n + 1)}{27 \cdot n^4} \geq 0. \tag{4.23}$$

which proves our claim. □

According to Lemma (5), the interception probability (4.15) is maximized for $p \in [0, \frac{1}{n}]$ at $\hat{p} = \frac{1}{n}$ with a maximum value of $a_3(n, \frac{1}{n}) = \frac{1}{n}$. This corresponds to a random walk with probability $\hat{r} = 0$ of remaining at the center. In particular, if the attack is initiated after any even number of periods with the Patroller being away from the attack node e, that is, with an even delay d, then the Patroller visits a single random end node in the three period attack interval, which means that he/she visits the attack node e with probability $\frac{1}{n}$, as seen by an alternative calculation above. We can summarize our findings into the following proposition;

Proposition 12. *The solution to the Uniformed Patroller Problem on the star network S_n with $m = 3$ is for the Patroller to follow a random walk with reflection at the ends (i.e. $r = 0$, $s = 1$), and for the Attacker to initiate her attack at an end node after $d = 2$ consecutive periods of the Patroller being away. Then, the value of the game (i.e. the probability that the optimal attack is intercepted) is given by $V = 1/n$.*

4.3.3 Attack duration $m = 4$

We finally consider the case where the attack lasts for four periods. Here, if the Attacker initiates her attack with the Patroller being at the center, then he gets two chances to intercept it, namely to find the correct end. However, if the attack starts with the Patroller being at an end node, then he gets only one chance. In the first case, the Patroller can intercept the attack with any of the following sequences:

$$ce_{--} \quad , \quad cce_{-} \quad , \quad ccce \quad , \quad ce^{'}ce,$$

where $e^{'}$ is any end node other than the attack end node e, while in the second case the Patroller can intercept the attack with either of the following two sequences:

$$e^{'}ce_{-} \quad , \quad e^{'}cce \quad , \quad e^{'}e^{'}ce.$$

It follows that the attack will be intercepted with overall probability:

$$\begin{aligned} a_4(n,p,s,q) &= q \cdot \big(1 + r + r^2 + (1-p-r)\cdot s\big)\cdot p \\ &\quad +(1-q)\cdot\big(s + s\cdot r + (1-s)\cdot s\big)\cdot p \\ &= q\cdot p\cdot\big(1 + r + r^2 + (1-p-r)\cdot s - s - s\cdot r - (1-s)\cdot s\big) \\ &\quad +p\cdot(2\cdot s + r\cdot s - s^2), \end{aligned} \tag{4.24}$$

where, like before, q is the conditional probability that the Patroller is at the center at the beginning of the attack given that he/she is not at the attack node e. Note that the coefficient of q in (4.24) is given by the product of p with the expression

$$\begin{aligned} & 1 + r\cdot(1-s) + r^2 + (1-p-r)\cdot s - s\cdot(2-s) \\ & \geq 1 + r\cdot(1-s) + r^2 + (1-p-r)\cdot s - 1 \\ & = r\cdot(1-s) + r^2 + (1-p-r)\cdot s \\ & \geq 0, \end{aligned} \tag{4.25}$$

since $r+p \leq r+n\cdot p = 1$, and $s \leq 1$. It follows that for fixed n and p, the interception probability (4.24) is increasing in q. Thus, by the same reasoning that we have used for $m = 2$ and $m = 3$, it further follows that the Attacker should choose to wait for $d = 2$ periods to attain $q = \hat{q}$, and the minimum interception probability of

$$a_4 = \frac{p}{1-p}\big((p+r-1)\cdot s^2 + 2\cdot(1-p-r\cdot p-r^2)\cdot s + r\cdot(1+r+r^2)\big). \tag{4.26}$$

To check that (4.26) is increasing in s, notice that the derivative with respect to s in the bracketed quadratic above is given by

$$\begin{aligned} & 2\cdot(p+r-1)\cdot s + 2\cdot(1-p-r\cdot p-r^2) \\ & \geq 2\cdot(p+r-1)\cdot 1 + 2\cdot(1-p-r\cdot p-r^2), \end{aligned} \tag{4.27}$$

since $r + p < 1$, which is equivalent to:

$$2 \cdot r \cdot (1 - p - r) \geq 0, \tag{4.28}$$

since both factors are non-negative.

Consequently, it turns out that the Patroller maximizes the interception probability for fixed p by choosing $s = 1$ (i.e. reflecting at the ends). Taking $s = 1$ in (4.26), gives:

$$\begin{aligned} a_4(n, p) &= \frac{p}{1-p} \cdot \left(2 \cdot r - p - 2 \cdot p \cdot r - r^2 + r^3 + 1\right) \\ &= \frac{-n^3 \cdot p^4 + 2 \cdot n^2 \cdot p^3 + 2 \cdot n \cdot p^3 - 3 \cdot n \cdot p^2 - 3 \cdot p^2 + 3 \cdot p}{1-p}. \end{aligned} \tag{4.29}$$

In Figure 4.5 we plot the variation of the interception probability $a_4(n, p)$ with p, for different number of nodes, while in Table 4.2 we give the maxima of $a_4(n, \hat{p})$ for the optimal $\hat{p}$ given by (4.7).

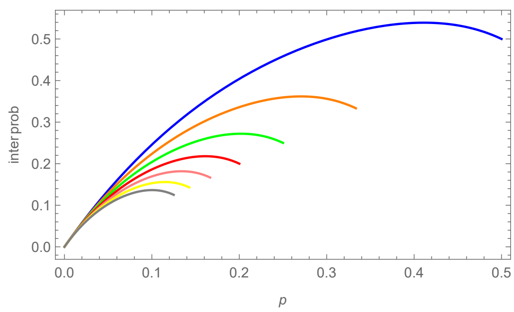


Figure 4.5: The interception probability $a_4(n, p)$, for $n = 2$ (blue), $\ldots, 8$ (grey).

n	$\hat{p}$	$a_4(n, \hat{p})$
2	0.4111	0.5391
3	0.2705	0.3618
4	0.2019	0.2720
5	0.1611	0.2179
6	0.1340	0.1817
7	0.1147	0.1559
8	0.1003	0.1364

Table 4.2: The interception probability $a_4(n, \hat{p})$, for $n = 2, \ldots, 8$.

4.4 The Line Network

The line network L_n consists of n nodes $i \in K$, and $n - 1$ edges. The two end nodes *1* and n are connected respectively to the penultimate nodes *2* and $n - 1$, while the $n - 2$ internal nodes $j \neq 1, n$, are connected respectively to nodes $j - 1$ to the left, and $j + 1$ to the right. We restrict the Patroller to Ergodic Markovian strategies, introducing the following setting. At the ends he/she reflects with probability κ, while he/she stays put with probability $\ell = 1 - \kappa$. From any other internal node $j \neq 1, n$, he/she moves towards the closest and the furthest end with probability p_j and q_j respectively, while he/she remains at the node with probability $r_j = 1 - p_j - q_j$.

To be consistent with this notation, we set $p_0 = p_n = 0$, $q_0 = q_n = \kappa$, and $r_0 = r_n = \ell$. In the special case when L_n consists of an odd number of nodes, the Patroller shifts from the center towards the two end nodes with equal probability c, staying around with probability $s = 1 - 2 \cdot c$.

We further assume that the Patroller is symmetric in his/her random patrolling, namely that:

$$p_j = p_{n+1-j} \quad \textit{and} \quad q_j = q_{n+1-j},$$

which ensures that the transition matrix A_n characterizing the Patroller's walk consists of $n - 1$ parameters.

This last assumption establishes a symmetry regarding the Attacker's strategy as well. That is, provided that the Patroller 'announces' an Ergodic Markovian patrol $(p_i, q_i, c, \kappa)_{i \in K}$, the Attacker anticipates equivalent interception probabilities (i.e. payoff) for her attack adopting either of the strategies (i, d) and $(n + 1 - i, d)$, where i, $n+1-i$ are symmetric, with respect to the center, nodes, and d is the number of periods the Attacker waits after the Patroller leaves her node before attacking.

We focus our attention on the analysis of an attack planned to take place at an end node, in particular at node *1*. The Patroller's initial position distribution in this case is $x^{(0)} = (1, 0, \ldots, 0)$. His/Her future position distribution conditional to him/her not having returned to the attack node *1* for t consecutive periods is denoted by $x^{(t)} = (x_1^{(t)}, \ldots, x_n^{(t)})$, where $x_1^{(t)} = 0$ for $t > 0$. Notice that whenever the Patroller returns to the attack node, t is reset to zero and his/her position distribution becomes again $x^{(0)}$. For that, we refer to $x^{(t)}$ as *the Patroller's away distribution* at $t > 0$. For example, the first two periods that the Patroller is away from node *1*, his/her away distribution is:

$$x^{(1)} = (0, 1, 0, \ldots, 0), \;\; x^{(2)} = (0, \frac{r_2}{1 - p_2}, \frac{q_2}{1 - p_2}, 0, \ldots, 0), \;\; \textit{etc.}$$

More generally, if we define the row vector $y^{(t)}$, for $t \geq 0$, by the product:

$$y^{(t)} = x^{(t)} \times A_n, \tag{4.30}$$

then the Patroller's away distribution at $t \geq 1$ is given by the following iteration formula:

$$x^{(t)} = \frac{x^{(t-1)} \times A_n - (y_1^{(t-1)}, 0, 0, 0)}{1 - y_1^{(t-1)}}. \tag{4.31}$$

Recall that we adopt a Stackelberg approach where the Patroller moves first, hence we set the Patroller to 'announce' first his/her strategy. Of course, in practice we interpret this as the Attacker observing the Patroller's motion on the network prior to deciding her strategy. The objective for the Patroller is to 'announce' the optimum strategy $(\hat{p}_i, \hat{q}_i, \hat{\kappa}, \hat{c})$ that maximizes the interception probability of an attack of known difficulty m for what is rationally the Attacker's optimum waiting time $\hat{t}$. Contrarily, the objective for the Attacker is to decide how many periods to wait in the absence of the Patroller from his/her chosen node before starting the attack. Namely, the Attacker aims to decide the optimum waiting time $\hat{t}$ that minimizes the interception probability of his/her attack for what is the announced-observed Markovian patrol. Hence, the interception probability of an attack under the two players' optimal strategies, that is the value V of the game, is given by:

$$V = \max_{A_n} \min_{t} \pi_m(p_i, q_i, \kappa, c, t) = \pi_m(\hat{p}_i, \hat{q}_i, \hat{\kappa}, \hat{c}, \hat{t}), \tag{4.32}$$

where $\pi_m(p_i, q_i, \kappa, c, t)$ is the interception probability of an attack at node *1* of duration m, under the patrol A_n and with attack delay $d = t$.

We solve this game numerically for the line networks L_4 and L_5, for given values of the attack duration m. The line network L_3 has been examined as a sub-case of the star network, since it is equivalent to S_2.

4.4.1 The Network L_4

First we analyse the uniformed patrolling game on the line network L_4 drawn below.

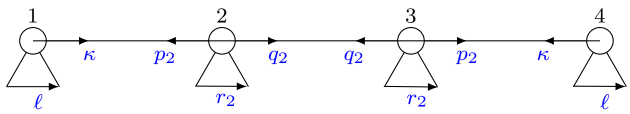


Figure 4.6: The line network L_4

To avoid unnecessary subscripts we introduce the following notation; $p_2 = p$, $q_2 = q$, $r_2 = r$. The general formula for recursively calculating the Patroller's away distribution, given by the vector equation (4.31), reduces to the following system:

$$\begin{aligned} (1 - p \cdot x_2^{(t-1)}) \cdot x_2^{(t)} &= (1 - p - q) \cdot x_2^{(t-1)} + q \cdot x_3^{(t-1)}, \\ (1 - p \cdot x_2^{(t-1)}) \cdot x_3^{(t)} &= (q - \kappa) \cdot x_2^{(t-1)} + (1 - p - q - \kappa) \cdot x_3^{(t-1)} + \kappa. \end{aligned} \tag{4.33}$$

For fixed p, q and κ, system (4.33) defines a continuously differentiable mapping $T_2 : \Delta^2 \to \Delta^2$. Any vector of the form $(0, x_2, x_3, x_4)$, such that $x_2, x_3, x_4 \in$

$(0,1)$ and $x_2 + x_3 + x_4 = 1$, will be a fixed point of T_2, if it also satisfies the system:

$$\begin{aligned} (p + q - p \cdot x_2) \cdot x_2 &= q \cdot x_3, \\ (\kappa + p + q - p \cdot x_2) \cdot x_3 &= \kappa + (q - \kappa) \cdot x_2. \end{aligned} \tag{4.34}$$

For the optimal patrolling strategies we estimate below, these equations have a unique solution which gives the limiting Patroller's away distribution under these optimal patrols. We aim to calculate the interception probability π_m of an attack at node *1*, considering five cases regarding its duration $m = 2, 3, 4, 5, 6$.

- For $m = 2$, if the Attacker initiates the attack when the Patroller is anywhere else but node *2*, he/she executes it without interception with probability 1. If however the Patroller is at node *2* when the attack starts, then it is intercepted with probability p. The interception probability is given by:

$$\pi_2(p, q, \kappa, t) = p \cdot x_2^{(t)}.$$

- For $m = 3$, if the Attacker initiates the attack when the Patroller is at the opposite end node, he/she executes it without interception with probability 1. If however the Patroller is at nodes *2* or *3* when the attack starts, then it is intercepted with non-zero probability. The overall interception probability is given by:

$$\pi_3(p, q, \kappa, t) = p \cdot \big(1 + (1 - p - q)\big) \cdot x_2^{(t)} + p \cdot q \cdot x_3^{(t)}.$$

- For $m = 4, 5, 6$ respectively, the Patroller intercepts an attack at node *1* with non-zero probability regardless of the node he/she is at when the Attacker initiates it. Indicatively, the overall interception probability for the attack duration $m = 4$ is given by:

$$\begin{aligned} \pi_4(p, q, \kappa, t) = p \cdot \big(1 + (1 - p - q) + (1 - p - q)^2 + q^2\big) \cdot x_2^{(t)} \\ + p \cdot q \cdot \big(1 + 2 \cdot (1 - p - q)\big) \cdot x_3^{(t)} + p \cdot q \cdot \kappa \cdot x_4^{(t)}, \end{aligned}$$

whereas we omit the lengthy explicit formulas for $\pi_5(p, q, \kappa, t)$ and $\pi_6(p, q, \kappa, t)$.

We have estimated numerically, for $d = t \leq 15$, the critical game values for the above five cases, and we present our findings in Table 4.3 rounded to four decimal places. Note that we include an additional column with the interception probabilities under the Patoller's optimal walk $(\hat{p}, \hat{q}, \hat{\kappa})$ for each case, but for the Attacker's waiting time $d = t \gg 15$. One can think of $\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \infty)$ as the interception probability the Attacker can expect if he/she waits a long time before starting his/her attack.

for $d = t \leq 15$	$(\hat{p}, \hat{q}, \hat{\kappa})$	$\hat{t}$	$\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \hat{t})$	$\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \infty)$
$m = 2$	$(0.3935, 0.3309, 1)$	4	0.1032	0.1067
$m = 3$	$(0.5, 0.5, 1)$	2	0.25	0.25
$m = 4$	$(0.4317, 0.4076, 1)$	4	0.2960	0.3158
$m = 5$	$(0.5, 0.5, 1)$	2	0.4375	0.4375
$m = 6$	$(0.4974, 0.4267, 1)$	4	0.4551	0.4766

Table 4.3: Critical Game Values for L_4 (Attack node *1*)

Remark 10. *We find that the Patroller should always reflect at the ends* ($\hat{\kappa} = 1$)*. Additionally, for an odd attack duration* ($m = 3, 5$) *the optimal patrol is a random walk* ($\hat{p} = \hat{q} = 0.5$) *that reflects at the boundaries, that is, the Patroller should never remain at the same node for two consecutive periods; respectively, the Attacker's best strategy is* $(1, 2)$*. For an even attack duration* ($m = 2, 4, 6$) *the Patroller optimally remains at an internal node with non-zero probability, which, however, decreases as we increase* m ($r_{m=2} > r_{m=4} > r_{m=6}$)*. We further observe that* $\hat{p} > \hat{q}$ *for an even* m*.*

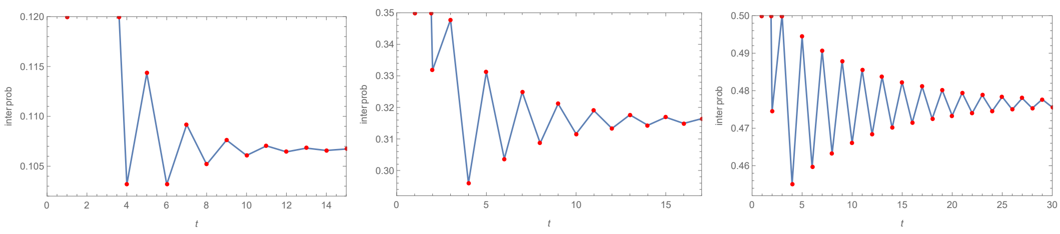
Indicatively, when $m = 4$ *the optimal values of* p *and* q *are respectively* 0.4317 *and* 0.4076*. The optimal Attacker's response to this Markovian patrol is to initiate his/her attack (at node 1) in the fourth period* ($d = 4$) *that the Patroller is away from the attack node. Under these optimal strategies the attack is intercepted with probability* 0.2960*. In the alternative case when the Attacker chooses to wait for a long time at the absence of the Patroller before attacking* ($d = t \gg 1$)*, his/her attack will be more likely to be intercepted (with probability* 0.3158*).*

However, there are two questions that naturally arise regarding our analysis, and need further investigation. To treat both of them we adopt a qualitative approach.

(*i*) *How confident we can be that for a waiting time* $d = t > 15$ *the Attacker cannot achieve a lower interception probability?*

A method to investigate this first issue is to generate the interception probabilities $\pi_m(p, q, \kappa, t)$ for increasing values of t (and for the optimum probabilities $\hat{p}, \hat{q}, \hat{\kappa}$ given in Table 4.3), and check whether we reach the limiting interception probabilities $\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \infty)$ without crossing below the optimum interception probabilities $\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \hat{t})$ that we have estimated for $d = t \leq 15$.

As we see in Figure 4.7, this appears to be the case for $m = 2$ (4.7a), $m = 4$ (4.7b), and $m = 6$ (4.7c), while for $m = 3$ and $m = 5$ we already reach the Patroller's stationary away distribution from the second period.



(a) $m = 2$ (b) $m = 4$ (c) $m = 6$

Figure 4.7: Interception probability for an attack at node *1* of L_4, under the optimal patrol, for increasing delay.

(ii) *Can the Attacker do any better by attacking a penultimate node* (*2 or 3*) *instead of attacking an end node?*

Regarding the second issue, we work as follows. For every attack duration m, we consider the optimum patrol $(\hat{p}, \hat{q}, \hat{\kappa})$ that we have estimated for an attack at the end node *1*, say for example the patrol $(0.4947, 0.4267, 1)$ for $m = 6$, and we generate the corresponding interception probability for an attack at node *2* for various waiting times $d = t$. Indicatively, for $m = 6$ we get

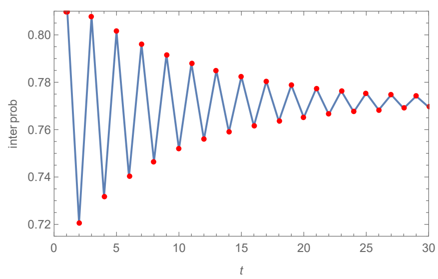


Figure 4.8: Interception probability for an attack at node *2* of L_4, under the optimal patrol estimated for an attack at node *1*, for increasing delay, for $m = 6$.

As we see in Figure 4.8, the minimum interception probability (0.7207) that we find for the attack strategy $(\mathit{2}, 1)$ is greater than the corresponding interception probability (0.4551) that we have found for the attack strategy $(\mathit{1}, 4)$, both estimated under the optimal patrol $(0.4974, 0.4267, 1)$, see Table 4.3. Hence, attacking at an end node, under the strategy $(\mathit{1}, 4)$, is the Attacker's optimal choice. Similar results we obtain for the rest of the cases, as we present in the following table.

for $(\hat{p}, \hat{q}, \hat{\kappa})$	$\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \hat{t})$/Node *1*	$\pi_m(\hat{p}, \hat{q}, \hat{\kappa}, \hat{t})$/Node *2*
$m = 2$	$(1, 4) - 0.1032$	$(2, 2) - 0.1363$
$m = 3$	$(1, 2) - 0.2500$	$(2, 1) - 0.5000$
$m = 4$	$(1, 4) - 0.2960$	$(2, 2) - 0.5237$
$m = 5$	$(1, 2) - 0.4375$	$(2, 1) - 0.7500$
$m = 6$	$(1, 4) - 0.4551$	$(2, 2) - 0.7207$

Table 4.4: Compare attacks at nodes *1* and *2* under the optimal patrol for node *1*

4.4.2 The Network L_5

Next we consider the line network L_5 which differs from L_4 in having a central node.

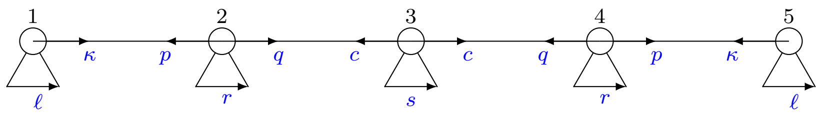


Figure 4.9: The line network L_5

Expanding the general recursive formula (4.31), we can rewrite the Patroller's away distribution for L_5 into the following system of equations:

$$\begin{aligned}
(1 - p \cdot x_2^{(t-1)}) \cdot x_2^{(t)} &= (1 - p - q) \cdot x_2^{(t-1)} + c \cdot x_3^{(t-1)}, \\
(1 - p \cdot x_2^{(t-1)}) \cdot x_3^{(t)} &= q \cdot x_2^{(t-1)} + (1 - 2 \cdot c) \cdot x_3^{(t-1)} + q \cdot x_4^{(t-1)}, \\
(1 - p \cdot x_2^{(t-1)}) \cdot x_4^{(t)} &= \kappa - \kappa \cdot x_2^{(t-1)} + (c - \kappa) \cdot x_3^{(t-1)} \\
&\quad + (1 - p - q - \kappa) \cdot x_4^{(t-1)}.
\end{aligned} \tag{4.35}$$

For fixed p, q, c and κ, system (4.35) defines a continuously differentiable mapping $T_3 : \Delta^3 \to \Delta^3$. Again, to be able to determine the limiting behavior of the Patroller's away distribution in L_5 for any fixed p, q, c and κ, we note that any stochastic vector $(0, x_2, x_3, x_4, x_5)$ is a fixed point of T_3, if it satisfies the system:

$$\begin{aligned}
(p + q - p \cdot x_2) \cdot x_2 &= c \cdot x_3, \\
(2 \cdot c - p \cdot x_2) \cdot x_3 &= q \cdot (x_2 + x_4), \\
(p + q + \kappa - p \cdot x_2) \cdot x_4 &= \kappa - \kappa \cdot x_2 + (c - \kappa) \cdot x_3.
\end{aligned} \tag{4.36}$$

For the optimal patrolling strategies we estimate below, these equations have a unique solution, that is the Patroller's away distribution the Attacker can expect if he/she waits a long time at the Patroller's absence before initiating the attack.

We can now determine the interception probability π_m corresponding to the Markovian parameters p, q, c and κ of the Patroller's walk, and the Attacker's delay $d = t$. We consider the same five cases regarding the attack duration, $m = 2, 3, 4, 5, 6$.

- In the first case, where $m = 2$, the Patroller may intercept the attack only by being at node *2* when it starts. Then, the interception probability is given by:

$$\pi_2(p, q, c, \kappa, t) = p \cdot x_2^{(t)}.$$

- In the second case, where $m = 3$, the Patroller intercepts the attack by being at either of nodes *2* or *3* when it starts. The overall interception probability:

$$\pi_3(p, q, c, \kappa, t) = p \cdot \big(1 + (1 - q - p)\big) \cdot x_2^{(t)} + p \cdot c \cdot x_3^{(t)}.$$

- In the third case, where $m = 4$, the Patroller intercepts the attack by being at nodes *2*, *3*, or *4* when it starts. The overall interception probability is given by:

$$\begin{aligned}\pi_4(p, q, c, \kappa, t) = p \cdot \big(1 + (1 - p - q) + q \cdot c + (1 - p - q)^2\big) \cdot x_2^{(t)} \\ + c \cdot p \cdot \big(1 + (1 - p - q) + (1 - 2 \cdot c)\big) \cdot x_3^{(t)} + q \cdot c \cdot p \cdot x_4^{(t)}.\end{aligned}$$

- In the last two cases, where $m = 5, 6$ respectively, the Patroller intercepts an attack at node *1* with non-zero probability regardless of the node he/she is at the beginning of the attack. We omit the explicit formulas for π_5 and π_6.

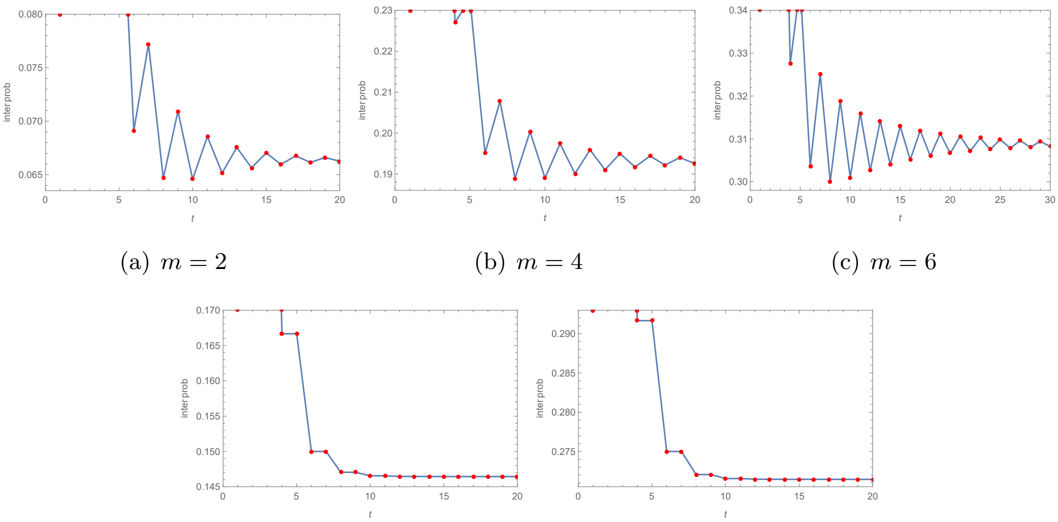
We have estimated numerically, for $d = t \leq 15$, the critical game values for the above five cases, and we present our findings in Table 4.5 rounded to four decimal places.

for $d = t \leq 15$	$(\hat{p}, \hat{q}, \hat{c}, \hat{\kappa})$	$\hat{t}$	$\pi_m(\hat{p}, \hat{q}, \hat{c}, \hat{\kappa}, \hat{t})$	$\pi_m(\hat{p}, \hat{q}, \hat{c}, \hat{\kappa}, \infty)$
$m = 2$	$(0.4342, 0.4110, 0.4096, 1)$	10	0.0646	0.0664
$m = 3$	$(0.5, 0.5, 0.5, 1)$	12	0.1464	0.1464
$m = 4$	$(0.4663, 0.4330, 0.4216, 1)$	8	0.1890	0.1932
$m = 5$	$(0.5, 0.5, 0.5, 1)$	14	0.2714	0.2714
$m = 6$	$(0.4821, 0.4598, 0.4332, 1)$	8	0.3	0.3088

Table 4.5: Critical game values for L_5 (Attack node *1*)

Remark 11. *As expected from the analysis for L_4, the Patroller should reflect at the ends in this case too. Similarly, we conclude that a random walk is the optimal patrol when the duration of the attack is odd (i.e. for $m = 3, 5$). A third observation is that for an even attack duration (i.e. for $m = 2, 4, 6$) the optimal $\hat{p}$ is always greater than the optimal $\hat{q}$, namely the optimal strategy for the Patroller is always to move towards the closest end with higher probability than move towards the center. One can conclude, as well, that increasing the attack duration for even values, the optimal patrol converges to a random walk, since $\hat{p}$, $\hat{q}$ and $\hat{c}$ increase in even m. Note that our findings are in principle identical with those obtained in Subsection 4.4.1 for the network L_4.*

However, the same issues arise here regarding the validity of our findings for a rather long delay $d = t > 15$, and the choice of the optimal attack node. To treat both we work the same way like before. Firstly, we generate the values of $\pi_m(\hat{p}, \hat{q}, \hat{c}, \hat{\kappa}, t)$, $m = 2, 3, 4, 5, 6$, for increasing values of t, and check whether we reach the limiting interception probabilities without crossing below the optimal interception probabilities that we have estimated for $d = t \leq 15$. As we see in Figure 4.10, this appears to be the case for every m, which validates our findings given in Table 4.5, for $d = t > 15$.

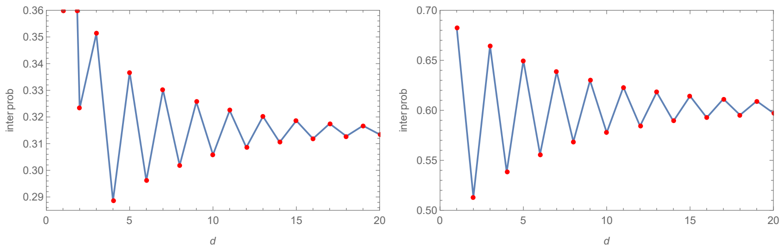


(a) $m = 2$ (b) $m = 4$ (c) $m = 6$

(d) $m = 3$ (e) $m = 5$

Figure 4.10: Interception probabilities for an attack at node *1* of L_5, under the optimal patrol, for increasing delay.

Regarding the optimal attack node, assuming the optimal patrols we have estimated for an attack at node *1*, first we generate the corresponding interception probabilities for an attack at nodes *2* and *3* for increasing delays t, then we identify their minima occurring at a specific t value, and last we compare those minima with the maxmin probabilities $\pi_m(\hat{p}, \hat{q}, \hat{c}, \hat{\kappa}, \hat{t})$ of Table 4.5. Indicatively, for $m = 4$ we obtain:



(a) Attack Node *2* (b) Attack Node *3*

Figure 4.11: Interception probability for an attack at nodes *2* and *3* of L_5, under the optimal patrol estimated for an attack at node *1*, for increasing delay, for $m = 4$.

As we see in Figures 4.11(a) and 4.11(b), for $m = 4$, the minimum interception probabilities 0.2886 and 0.5132, that we find respectively for the attack strategies $(2, 4)$ and $(3, 2)$, are both greater than the minimum interception probability 0.1890 that we have found for the attack strategy $(1, 8)$ (all estimated for the optimal patrol $(0.4663, 0.4330, 0.4216, 1)$, see Table 4.5). Similar results, confirming that the Attacker should comparatively attack at an end node (node *1* in this instance), we obtain for the rest of the m values we have considered, as we present in Table 4.6.

for $(\hat{p}, \hat{q}, \hat{c}, \hat{\kappa})$	π_m/Node 1	π_m/Node 2	π_m/Node 3
$m = 2$	$0.0646 - (1, 10)$	$0.0876 - (2, 4)$	$0.1080 - (3, 2)$
$m = 3$	$0.1464 - (1, 12)$	$0.25 - (2, 2)$	$0.5 - (3, 1)$
$m = 4$	$0.1890 - (1, 8)$	$0.2886 - (2, 4)$	$0.5132 - (3, 2)$
$m = 5$	$0.2714 - (1, 14)$	$0.4375 - (2, 2)$	$0.75 - (3, 1)$
$m = 6$	$0.3 - (1, 8)$	$0.4545 - (2, 4)$	$0.7473 - (3, 2)$

Table 4.6: Compare attacks at nodes *1* and *2*, *3* under the optimal patrol for node *1*

4.5 The Circle Network

The circle network C_n is a closed walk consisting of n nodes, and equal number of edges. Like above, we restrict the Patroller to Ergodic Markovian strategies such that the Patroller leaves every node with overall probability $2 \cdot p$, moving clockwise and anticlockwise with the same probability p, while he/she remains at the same node with probability $r = 1 - 2 \cdot p$. Given this symmetric patrol, every strategy (i, d), for a random node i and a given delay d, is equally advantageous for the Attacker.

According to our setting, the transition matrix B_n characterizing the Patroller's Ergodic Markovian walk on the circle network is parametrized by a single parameter p. Without loss of generality, we take node *1* as the attack node.

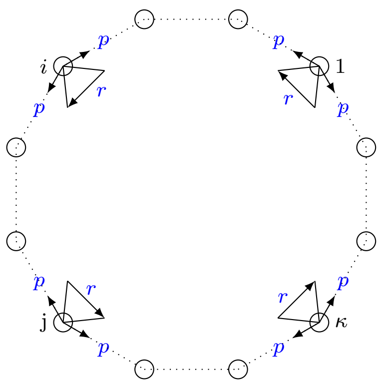


Figure 4.12: The circle network C_n

We denote the *Patroller's away distribution from node 1*, for $t > 0$, by $x^{(t)} = (0, x_2^{(t)}, \dots, x_n^{(t)})$, where $x^{(0)} = \{1, 0, \dots, 0\}$. Equivalently to our approach in Section 4.4, if we define the row vector $y^{(t)}$, for $t \geq 0$, by:

$$y^{(t)} = x^{(t)} \times B_n, \tag{4.37}$$

then the Patroller's away distribution from node *1*, at $t \geq 1$, is given by the iteration:

$$x^{(t)} = \frac{x^{(t-1)} \times B_n - (y_1^{(t-1)} 0, 0, 0)}{1 - y_1^{(t-1)}}. \tag{4.38}$$

We consider networks C_4, C_5, and we solve the game numerically for given values of m.

4.5.1 The Network C_4

We start with the circle network with four nodes. Equivalently to (4.38), we can rewrite the Patroller's away distribution for C_4 into the following system of equations:

$$\begin{aligned} (1 - p + p \cdot x_3^{(t-1)}) \cdot x_2^{(t)} &= (1 - 2 \cdot p) \cdot x_2^{(t-1)} + p \cdot x_3^{(t-1)}, \\ (1 - p + p \cdot x_3^{(t-1)}) \cdot x_3^{(t)} &= p + (1 - 3 \cdot p) \cdot x_3^{(t-1)}, \end{aligned} \tag{4.39}$$

which defines a continuously differentiable map $T(x_2^{(t)}, x_3^{(t)}) = (x_2^{(t+1)}, x_3^{(t+1)})$. There is a unique stationary solution of T of the form $(0, x_2, x_3, x_4)$, independent of p, given by:

$$(0, x_2, x_3, x_4) = (0, \frac{2-\sqrt{2}}{2}, \sqrt{2}-1, \frac{2-\sqrt{2}}{2}). \tag{4.40}$$

We take four cases regarding the attack duration, $m = 2, 3, 4, 5$. Indicatively we have:

- In the first case, where $m = 2$, the interception probability is given by:

$$\pi_2(p, t) = p \cdot (x_2^{(t)} + x_4^{(t)}).$$

- In the second case, where $m = 3$, the interception probability is given by:

$$\pi_3(p, t) = 2 \cdot p \cdot (1-p) \cdot (x_2^{(t)} + x_4^{(t)}) + 2 \cdot p^2 \cdot x_3^{(t)}.$$

- In the third case, where $m = 4$, the interception probability is given by:

$$\begin{aligned}\pi_4(p, t) = p \cdot \big(2 - 2 \cdot p + (1 - 2 \cdot p)^2 + 2 \cdot p^2\big) \cdot (x_2^{(t)} + x_4^{(t)}) \\ +2 \cdot p^2 \cdot (3 - 4 \cdot p) \cdot x_3^{(t)}.\end{aligned}$$

- In the fourth case, where $m = 5$, the interception probability is given by:

$$\begin{aligned}\pi_5(p, t) = p \cdot \big(2 - 2 \cdot p + (1 - 2 \cdot p)^2 + (1 - 2 \cdot p)^3 + 2 \cdot p^2 \cdot (4 - 6 \cdot p)\big) \\ \times (x_2^{(t)} + x_4^{(t)}) + 2 \cdot p^2 \cdot \big(2 \cdot p^2 + 2 \cdot (1 - 2 \cdot p)^2 + 3 - 4 \cdot p\big) \cdot x_3^{(t)}.\end{aligned}$$

We have estimated numerically, for $d = t \leq 15$, the critical game values for the above four cases, and we present our results in Table 4.7 rounded to four decimal places.

for $d = t \leq 15$	$\hat{p}$	$\hat{t}$	$\pi_m(\hat{p}, \hat{t})$	$\pi_m(\hat{p}, \infty)$
$m = 2$	0.2929	2	0.1716	0.1716
$m = 3$	0.5	1	0.5	0.5
$m = 4$	0.4515	2	0.5216	0.6021
$m = 5$	0.5	1	0.75	0.75

Table 4.7: Critical game values for C_4

Like before, we aim to examine (qualitatively) whether our results are valid for a relatively long delay $d = t > 15$. For $m = 2, 3, 5$, the optimal interception probabilities coincide with the corresponding limiting interception probabilities (see Table 4.7).

Accordingly, for $m = 4$, we reach the limiting interception probability without crossing below the optimum interception probability that we have estimated for $d = t \leq 15$ (see Figure 4.13). Thus, we claim our results are valid for infinitely large delay $d = t$.

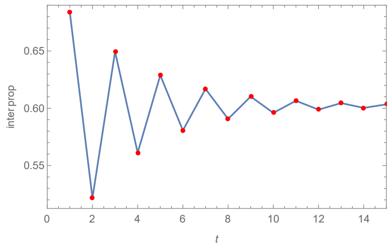


Figure 4.13: Interception probability for an attack at a random node of C_4 under the optimal patrol, for increasing delay, for $m = 4$.

4.5.2 The Network C_5

Next we consider the circle network with five nodes. Equivalently to (4.38), we can rewrite the Patroller's away distribution from node *1* into the following system:

$$\begin{aligned} (1-p+p\cdot x_3^{(t-1)}+p\cdot x_4^{(t-1)})\cdot x_2^{(t)} &= (1-2\cdot p)\cdot x_2^{(t-1)}+p\cdot x_3^{(t-1)}, \\ (1-p+p\cdot x_3^{(t-1)}+p\cdot x_4^{(t-1)})\cdot x_4^{(t)} &= p+(1-3\cdot p)\cdot x_4^{(t-1)}-p\cdot x_2^{(t-1)}, \\ (1-p+p\cdot x_3^{(t-1)}+p\cdot x_4^{(t-1)})\cdot x_3^{(t)} &= p\cdot(x_2^{(t-1)}+x_4^{(t-1)}) \\ &\quad +(1-2\cdot p)\cdot x_3^{(t-1)}, \end{aligned} \tag{4.41}$$

which defines a continuously differentiable mapping $T_3 : \Delta^3 \to \Delta^3$. There is a unique stationary solution of T_3 of the form $(0, x_2, x_3, x_4, x_5)$, that is also independent of p:

$$(0, x_2, x_3, x_4, x_5) = (0, \frac{3-\sqrt{5}}{4}, \frac{\sqrt{5}-1}{4}, \frac{\sqrt{5}-1}{4}, \frac{3-\sqrt{5}}{4}). \tag{4.42}$$

We take here the same four cases as in Section 4.5.1 regarding the attack duration m.

- In the first case, where $m = 2$, the interception probability is given by:

$$\pi_2(p,t) = p\cdot(x_2^{(t)}+x_5^{(t)}),$$

- In the second case, where $m = 3$, the interception probability is given by:

$$\pi_3(p,t) = 2\cdot p\cdot(1-p)\cdot(x_2^{(t)}+x_5^{(t)})+p^2\cdot(x_3^{(t)}+x_4^{(t)}),$$

- In the third case, where $m = 4$, the interception probability is given by:

$$\pi_4(p,t) = p \cdot \left(2 - 2 \cdot p + (1 - 2 \cdot p)^2 + p^2\right) \cdot (x_2^{(t)} + x_5^{(t)}) \\ +3 \cdot p^2 \cdot (1 - p) \cdot (x_3^{(t)} + x_4^{(t)}),$$

- In the last case, where $m = 5$, the interception probability is given by:

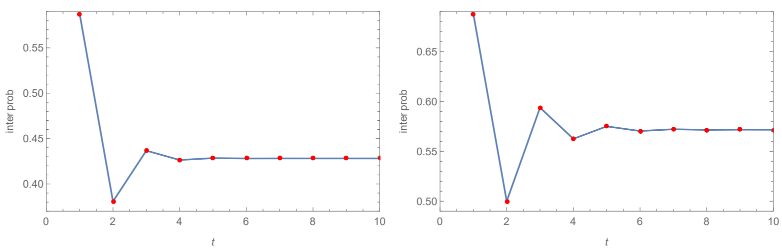
$$\pi_5(p,t) = p \cdot \left(2 - 2 \cdot p + (1 - 2 \cdot p)^2 + p^2 + 2 \cdot p^2 \cdot (1 - 2 \cdot p) + p^3\right) \\ \times (x_2^{(t)} + x_5^{(t)}) + p^2 \cdot \left(3 - 4 \cdot p^2 + 3 \cdot (1 - 2 \cdot p)^2\right) \cdot (x_3^{(t)} + x_4^{(t)}).$$

We have estimated numerically, for $d = t \leq 15$, the critical game values for the above four cases, and we gather our results in Table 4.8 rounded to four decimal places.

for $d = t \leq 15$	$\hat{p}$	$\hat{t}$	$\pi_m(\hat{p}, \hat{t})$	$\pi_m(\hat{p}, \infty)$
$m = 2$	0.3820	2	0.1459	0.1459
$m = 3$	0.3820	2	0.2705	0.2705
$m = 4$	0.4450	2	0.3808	0.4282
$m = 5$	0.5	2	0.5	0.5716

Table 4.8: Optimal game values for C_5

Like before, we want to examine whether increasing the attack duration $d = t \geq 15$ (for constant $p = \hat{p}$), we reach the limiting interception probability without crossing below the optimum interception probability that we have estimated for $d = t \leq 15$. For $m = 4, 5$ this appears to be the case as we see in Figure 4.14, while for $m = 2, 3$ the optimum interception probability coincides with the limiting interception probability (see Table 4.8).



(a) $m = 4$ (b) $m = 5$

Figure 4.14: Interception probability for an attack at a random node of C_5 under the optimal patrol, for increasing delay.

4.6 The Star in the Circle Network

The star in the circle network E_n combines the star and the circle networks we have considered above. It consists of n end nodes, one central node, and $2 \cdot n$ edges. As usual, we take a Markovian Patroller, leaving an end node towards an adjacent end node with probability p, and towards the centre with probability q, moves from the centre towards each of the end nodes with probability r, while remains at an end node with probability $a = 1 - 2 \cdot p - q$, and at the centre with probability $b = 1 - n \cdot r$.

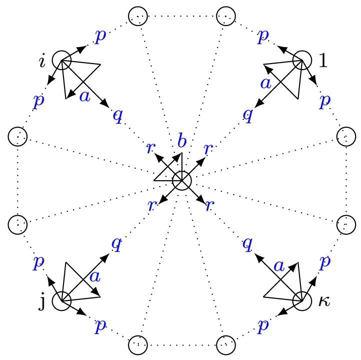


Figure 4.15: The star in the circle network E_n

The transition matrix Γ_n describing the Patroller's walk on the star in the circle network is parametrized by three parameters, p, q, r. Regarding the Attacker, he/she has two options. Either to attack, equally preferably, at one of the ends, or to attack at the centre. To highlight the Attacker's bounded rationality, here we introduce an additional strategic element. We set the Attacker holding a counter of limited storing capacity counting his/her delay. When the Patroller leaves the attack node, the Attacker turns on the counter recording periods at the Patroller's absence up to a finite level D (max delay). If the Patroller has not returned after D periods, the Attacker can no longer count his/her absence and adjust the attack, that is, the Attacker can rationally initiate the attack after waiting at most for D periods. We call such an Attacker a *Finite Automata Agent*, in alignment with Rubinstein [1986]. Without loss of generality, we take the end node *1* as the attack node. In accordance with (4.31) and (4.44), the Patroller's away distribution from node *1*, for $t \geq 1$, is:

$$x^{(t)} = \frac{x^{(t-1)} \times \Gamma_n - (z_1^{(t-1)}, 0, 0, 0)}{1 - z_1^{(t-1)}}, \tag{4.43}$$

where $z^{(t)} = x^{(t)} \times \Gamma_n$ is the row vector defined for $t \geq 0$, and $x^{(0)} = (1, 0, \ldots, 0)$.

The Patroller's away distribution from the centre, for $t \geq 1$, is defined by

$y^{(t)} = (y_1^{(t)}, \dots, y_n^{(t)}, 0)$. For an attack at the centre of duration m, the interception probability is given by:

$$\pi_c(q, m) = (m-1) \cdot q. \tag{4.44}$$

Recall that we aim to find the optimal patrol $(\hat{p}, \hat{q}, \hat{r})$, and the optimal delay $\hat{t}$. First though we need to examine which node the Attacker optimally attacks. Obviously, if the Patroller announces $\hat{q}$ such that $\pi_c(\hat{q}, m) < \pi_m(p, \hat{q}, r, t)$ then the Attacker should attack the centre. Therefore, here we deal with the following problem:

$$\max_{\Gamma_n} \min_t \pi_m(p, q, r, t) = \pi_m(\hat{p}, \hat{q}, \hat{r}, \hat{t}) : \pi_m(\hat{p}, \hat{q}, \hat{r}, \hat{t}) \leq \pi_c(\hat{q}, m), \tag{4.45}$$

the solution of which actually makes the Attacker indifferent of which node to attack.

4.6.1 The Network E_3

First we consider the star in the circle network with three ends. The interception probability of an attack at node *1* of duration $m = 2, 3, 4$, as a function of p, q, r, t, is:

- for $m = 2$,
$$\pi_2(p, q, r, t) = p \cdot (x_2^{(t)} + x_3^{(t)}) + r \cdot x_c^{(t)},$$
- for $m = 3$,
$$\pi_3(p, q, r, t) = (p + a \cdot p + p^2 + q \cdot r) \cdot (x_2^{(t)} + x_3^{(t)}) + (r + 2 \cdot r \cdot p + b \cdot r) \cdot x_c^{(t)},$$
- for $m = 4$,
$$\pi_4(p, q, r, t) = r \cdot (2 \cdot p + b + 1 + b^2 + 2 \cdot q \cdot r + p^2 + 2 \cdot b \cdot p + 2 \cdot a \cdot p) \cdot x_c^{(t)}$$
$$+ \big(p \cdot (1 + a + p + a^2 + 2 \cdot a \cdot p + 3 \cdot q \cdot r + p^2) + q \cdot r \cdot (1 + a + b)\big) \cdot (x_2^{(t)} + x_3^{(t)}).$$

We solve the game numerically, for $D = 15$, and we gather our results in Table 4.9, rounded to four decimal places.

for $D = 15$	$(\hat{p}, \hat{q}, \hat{r})$	$d = \hat{t}$	$\pi_m(\hat{p}, \hat{q}, \hat{r}, \hat{t})$
$m = 2$	$(0.3333, 0.3333, 0.3333)$	1	0.3333
$m = 3$	$(0.3560, 0.2881, 0.3333)$	2	0.5761
$m = 4$	$(0.3762, 0.2476, 0.3333)$	2	0.7428

Table 4.9: Critical game values for E_3

4.6.2 The Network E_4

Next we consider the star in the circle network with four end nodes. The interception probability of an attack at node *1* of duration $m = 2, 3, 4$ as a function of p, q, r, t is:

- for $m = 2$,
$$\pi_2(p,q,r,t) = p \cdot (x_2^{(t)} + x_4^{(t)}) + r \cdot x_c^{(t)},$$

- for $m = 3$,
$$\pi_3(p,q,r,t) = (p + a \cdot p + q \cdot r) \cdot (x_2^{(t)} + x_4^{(t)}) + (2 \cdot p^2 + q \cdot r) \cdot x_3^{(t)} \\ +r \cdot (1 + 2 \cdot p + b) \cdot x_c^{(t)},$$

- for $m = 4$,
$$\pi_4(p,q,r,t) = r \cdot (1 + 2 \cdot p + b + 2 \cdot b \cdot p + 2 \cdot a \cdot p + b^2 + 3 \cdot r \cdot q + 2 \cdot p^2) \cdot x_c^{(t)} \\ +\big(2 \cdot p^2 + q \cdot r + 4 \cdot a \cdot p^2 + 4 \cdot p \cdot q \cdot r + a \cdot q \cdot r + q \cdot b \cdot r\big) \cdot x_3^{(t)} \\ +\big(2 \cdot p^3 + 3 \cdot q \cdot r \cdot p + p + a \cdot r + q \cdot r + a^2 \cdot p + a \cdot q \cdot r + q \cdot b \cdot r\big) \cdot (x_2^{(t)} + x_4^{(t)}).$$

We solve the game numerically, for $D = 15$, and we gather our results in Table 4.10, rounded to four decimal places.

for $D = 15$	$(\hat{p}, \hat{q}, \hat{r})$	$d = \hat{t}$	$\hat{\pi}_m(\hat{p}, \hat{q}, \hat{r}, \hat{t})$
$m = 2$	$(0.2835, 0.1695, 0.25)$	2	0.1695
$m = 3$	$(0.3993, 0.2013, 0.25)$	2	0.4026
$m = 4$	$(0.4147, 0.1706, 0.25)$	2	0.5118

Table 4.10: Critical game values for E_4

4.7 Variations with a Non-Markovian Patroller

In this last section, we drop the requirement that the Patroller adopts a Markovian strategy, allowing him/her to adopt arbitrary walks on the patrolling network Q. By extension, we relate variations of the *Uniformed Patroller problem* introduced here, with two deterministic problems on networks considered extensively in the computer science literature. We take the search-patrolling region Q to be a finite network with integer edge lengths, that is, with integer travel times.

4.7.1 The UAV-Pinwheel Problem

Suppose a reconnaissance unmanned aerial vehicle (UAV) must return to locations i at most every κ_i periods. Is this possible when the locations i are the nodes of a given network Q, where $d(i,j)$ is the given integer number of periods to travel between nodes i and j? If not, how many UAV's K are required to make this task possible? This problem is called the *cyclic routing UAV (CRUAV) problem*, see, e.g., Ho and Ouaknine [2015]. For $K = 1$ it is related to a winning condition for the Patroller in an extended patrolling game $U^* = U^*(Q, m_1, m_2, \ldots, m_n)$, where m_i is the duration of an attack at node i of Q (in our original setting $m_i = m$ for every node i). If there is a solution to the CRUAV problem with $\kappa_i = m_i - 1$, then this solution is a winning solution for the Patroller in the patrolling game U^*. Of course such strategies are non-Markovian.

To illustrate this problem, consider the network drawn below as taken from Ho and Ouaknine [2015]. This problem has a positive solution with $K = 1$, that is, the periodic walk $DCABACDACBA$, which has period 11. The return times to A are $2, 5, 5, 5$ (≤ 5), to B are $7, 10$ (≤ 10), to C are $5, 5, 6$ (≤ 6), and to D are $8, 9$ (≤ 9). Note that no walk of period 4 (a minimum period walk visiting all nodes) returns to node A (node with the minimum attack duration) in time 5, as the required edges have at most two with length 1 and the rest with length 2, that is, such a four period walk/patrol lasts at least for $1 + 1 + 2 + 2 = 6$ periods. Hence, the Uniformed Patroller problem on this network has value 1, that is, the Patroller intercepts any attack of the given durations with probability 1.

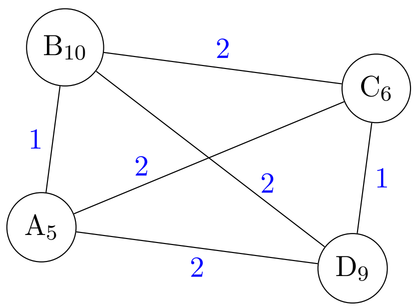


Figure 4.16: Example of the cyclic routing UAV problem. Ho and Ouaknine [2015]

Apparently, the CRUAV problem is a variant of the so-called *Pinwheel Problem*, see, e.g., Holte et al. [1989], and of the *Patrolling Security Games*, as those have been discussed in Basilico, Gatti and Amigoni [2012], Basilico, De Nittis and Gatti [2015]. In principle, the CRUAV problem has direct, important connections with the general family of scheduling problems.

4.7.2 Spy Games

Our model gives the Attacker additional information compared with the original Patrolling Games, that is, he/she can see whether the Patroller is located at the same location with him/her. Suppose that we let the Attacker can see the Patroller wherever he/she is. We keep our original assumption that the attacks at any node require the same number of periods m. In this case the Attacker may want to move, for example in a direction away from the Patroller, before starting his/her attack. We allow the two players to alternate in moving, with the Patroller constrained to moving only to an adjacent node per turn, and the Attacker constrained to moving to a node at a distance at most S nodes away per turn. Of course, during his/her attack the Attacker must stay still. In the 'spy-game' introduced by Cohen et al. [2016], the Attacker (identified as the Robber) wins if he/she ever gets a distance δ away from the Patroller (identified as the Cop, who in their game can see the Attacker). If we take $\delta = m + 1$, and tell the Attacker to begin his/her attack (and stay still ever since) as soon as he/she has achieved distance δ, then the Patroller cannot intercept the attack given his/her unit speed, their initial distance and the attack duration. This idea yields a connection between the Patroller-Attacker games and what are known as the Cops and Robber games, see, Aigner and Fromme [1984].

Chapter 5

Conclusion

In Chapter 2, we study the distribution of illegal activity in a population of myopic inspectees (e.g., tax payers) interacting under the pressure of a short-sighted benevolent inspector (e.g., tax inspector). Equivalently, we investigate the spread of corruption in a population of myopic bureaucrats (e.g., ministerial employees) interacting under the pressure of an incorruptible supervisor(e.g., governmental fraud investigator). We consider two game settings with regards to the inspectees' available strategies, where the continuous strategy setting is a natural extension of the discrete strategy setting we initially consider. We introduce, and vary both qualitatively and quantitatively, two key control elements that govern the deterministic evolution of the illegal activity in the population, the punishment fine and the inspection budget.

We derive the ODEs (2.12) and (2.27) that characterize the dynamics of our system. In particular, we identify explicitly the fixed points that occur under our different scenarios, and we carry out respectively their stability analysis. We show that although the indistinguishable 'agents' (e.g., inspectees, corrupted bureaucrats) are treated as myopic maximizers, profitable strategies eventually prevail in their population through imitation. We verify that an adequately financed inspector achieves an increasingly law-abiding environment when there is a stricter fine policy. We show, however, that although the inspector can establish any desirable average violation by suitably manipulating his/her renewable budget and his/her fine policy, he/she is not able to combine this with any desirable group's strategy profile. Finally, we provide a game-theoretic interpretation of the limiting dynamics fixed points stability analysis (link with the Nash equilibria of the corresponding N-player game).

There are many directions towards which one can extend our approach. To begin with, one can consider an inspector experiencing policy-adjusting costs. Equivalently, one can withdraw the assumption of a renewable budget. Moreover, the case of two or more inspectors, possibly collaborating with each other or even with some of the inspectees, could be examined. Regarding the inspectees, an additional source of interaction based on the social norms established within their population could be introduced. Another interesting variation would be to add a spatial distribution of the population, assuming indicatively that the inspectees interact on a specific network. Some of these alternatives have been studied for similar game settings; see, e.g., Kolokoltsov [2014], Kolokoltsov and Malafeyev [2015].

In Chapter 3 we formulate the interaction of a large number of small players under the pressure of a major player (principal), on n-dimensional arrays, having in mind the paradigm of individuals defending against a bio-terrorist; alternatively, the similar context of corrupted tax inspectors against a benevolent authority. The n-dimensional arrays dual structure naturally describes on the one hand the distribution of individuals among m levels of 'behaviour' (e.g. levels of defence) and on the other, their distribution according to a phenotypic characteristic among n levels of 'hierarchy' (e.g. levels of infection). Transitions on the first structure is mainly subject to the individuals' control, while transitions on the second is mainly subject to the principal's pressure. Transitions on both structures may also be an outcome of the individuals' binary interactions. Our model is a performance of a finite state non-linear Markov game combining mean-field, evolutionary, and pressure-resistance types of interaction.

For our analysis, we consider the discounted mean-field game consistency problem. We demonstrate the kinetic equations governing the evolution of the individuals' distribution among the $n \times m$ states (forward equation), and the Hamilton-Jacobi-Bellman equation giving the individuals' optimal payoff (backward equation). We solve the stationary problem and we provide a link of the stationary solution to the time-dependent problem. For simplicity, we work in the asymptotic regimes of fast execution of personal decisions, weak binary interactions, and small payoff discounting in time. Considering a stationary control that is consistent with the assumption of fast execution of personal decisions, in the main order of small payoff discounting in time (or in the main order of weak binary interactions), we find that individuals will be uniformly distributed among the 'behaviours' of the unique 'hierarchy' level where the sum of rewards is maximised, and we obtain the optimal payoff as a function of these rewards. We show that there is a unique solution to the time-dependent problem, that is very close to the stationary solution.

Our simplifications, while necessary for concrete calculations, consist only the first step towards a more comprehensive treatment of the game we have formulated.

In Chapter 4, we introduce the characteristic of limited vision to the Attacker in a network patrolling game by letting him/her see, at the beginning of every period, whether or not the Patroller is present at the node he/she is planning to attack. Obviously, the Attacker will not initiate the attack when the Patroller is present, and we further show here that optimally he/she will also not immediately attack when the Patroller leaves his/her node, but will rather wait an optimal number of periods, resetting his/her count if the Patroller returns before that time. This behaviour is in stark contrast to the optimal behaviour when the Patroller follows a known periodic tour of the network, where an immediate attack is often the best strategy. For example many prison escape movies show the prisoners (i.e. the Attackers) attempting to escape just after the spotlight (i.e. the Patroller) leaves their location.

We have adopted here the assumption that the Patroller's motion can be observed (and therefore be recorded) prior to the beginning of the game, or generally speaking, that is somehow a known element of the game. That is, we consider a Stackelberg game approach. However, an interesting observation of our analysis and results for several networks is that, in fact, the Attacker has an optimal strategy that does not require any prior knowledge of the Patroller's motion. In other words, there exists a Nash equilibrium between the Patroller and the Attacker.

A number of possible extensions of our model naturally suggest themselves after our results are further taken into consideration. For one, it would be interesting to give the Attacker a greater range of vision, perhaps initially to nodes adjacent to the planned attack node. Of course, if the Attacker has significant vision in this respect, he/she might have an incentive to move away from the Patroller's current location by changing his/her choice of which node to attack. This approach might then lead to one-sided information in the so called Cops and Robbers games, with only the Robbers (the Attacker in our case) having vision. In terms of the search and pursuit-evasion games wording, this might be called a search-evasion game approach. Another extension to our model, seemingly rather difficult to deal with, would be to provide the Patroller with the same range of motions, that is mixtures over general walks, that was allowed in the original formulation of network patrolling games, Alpern, Morton and Papadaki [2011]. A first step towards this direction would be to introduce Markovian strategies with short memories.

Moreover, our model raises the question of why patrols are ever carried out with uniforms that give to a potential attacker or infiltrator additional helpful information. One answer might lie in the direction of deterrence. Under some of our game parameters, where the interception probability is relatively high, the Attacker might choose to abandon his/her attempt altogether, leading to a non zero-sum game. Clearly, the Uniformed Patroller model leads to many open, unanswered questions.

Bibliography

Aidt T S (2003) Economic analysis of corruption: A survey. *The Economic Journal*. 113(491):632-652.

Aigner M, Fromme M (1984) A game of cops and robbers. *Discrete Applied Mathematics*. 8(1):1-12.

Alferov G V, Malafeyev O A, Maltseva A S (2015) Programming the robot in tasks of inspection and interception. In Mechanics-Seventh Polyakhov's Reading, 2015 International Conference on (pp. 1-3). IEEE.

Alpern S, Morton A, Papadaki K (2011) Patrolling games. *Operations Research*. 59(5):1246-1257.

Arce M, Daniel G, Sandler T (2005) Counterterrorism: A game-theoretic analysis. *Journal of conflict resolution*. 49(2):183-200.

Aumann R J (1964) Markets with a continuum of traders. *Econometrica: Journal of the Econometric Society*. 39-50.

Avenhaus R, von Stengel B (1992) Non-zero-sum Dresher inspection games. *Operations Research*. 91:376-379

Avenhaus R, Canty M, Kilgour D M, von Stengel B, Zamir S (1996) Inspection games in arms control. *European Journal of Operational Research*. 90(3):383-394.

Avenhaus R, von Stengel B, Zamir S (2002) Inspection games. *Handbook of game theory with economic applications*. 3:1947-1987.

Avenhaus R (2004) Applications of inspection games. *Mathematical Modelling and Analysis*. 9(3):179-192.

Avenhaus R, Kilgour D M (2004) Efficient distributions of arms-control inspection effort. *Naval Research Logistics (NRL)*. 51(1):1-27.

Avenhaus R, Canty M J (2012) Inspection games. In Computational Complexity: Theory, Techniques, and Applications. (pp. 1605-1618) Springer New York.

Basilico N, Gatti N, Amigoni F (2012) Patrolling security games: Definition and algorithms for solving large instances with single patroller and single intruder. *Artificial Intelligence.* 184:78-123.

Basilico N, De Nittis G, Gatti N (2015) Adversarial patrolling with spatially uncertain alarm signals. arXiv preprint arXiv:1506.02850.

Basna R, Hilbert A, Kolokoltsov V N (2014) An epsilon-Nash equilibrium for non-linear Markov games of mean-field-type on finite spaces. *Communications on Stochastic Analysis.* 8(4), 3.

Baston V J, Bostock F A (1991) A generalized inspection game. *Naval Research Logistics (NRL).* 38(2):171-182

Baston V, Kikuta K (2004) An ambush game with an unknown number of infiltrators. *Operations Research.* 52(4):597-605.

Baston V, Kikuta K (2009) Technical Note-An Ambush Game with a Fat Infiltrator. *Operations Research.* 57(2):514-519.

Baykal-Gürsoy M, Duan Z, Poor H V, Garnaev A (2014) Infrastructure security games. *European Journal of Operational Research.* 239(2):469-478.

Bensoussan A, Frehse J, Yam P (2013) *Mean field games and mean field type control theory.* New York: Springer.

Bensoussan A, Chau M H M, Yam S C P (2016) Mean field games with a dominating player. *Applied Mathematics & Optimization.* 74(1):91-128.

Bergin J, Bernhardt D (1992) Anonymous sequential games with aggregate uncertainty. *Journal of Mathematical Economics.* 21(6):543-562.

Caines P E (2013) Mean field games. *Encyclopedia of Systems and Control.* 1-6.

Canty M J, Rothenstein D, Avenhaus R (2001) Timely inspection and deterrence. *European Journal of Operational Research.* 131(1):208-223.

Cardaliaguet P (2010) Notes on mean field games (p. 120). Technical report.

Carmona R, Delarue F (2013) Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization.* 51(4):2705-2734.

Carmona R, Zhu X (2016) A probabilistic approach to mean field games with major and minor players. *The Annals of Applied Probability.* 26(3):1535-1580.

Cohen N, Hilaire M, Martins N, Nisse N, Prennes S (2016) Spy-Game on graphs. In 8th International Conference on Fun with Algorithms, FUN 2016 (Vol. 49). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Collins A, Czyzowicz J, Gasieniec L, Kosowski A, Kranakis E, Krizanc D, ..., Morales Ponce O (2013, July). Optimal patrolling of fragmented boundaries. In Proceedings of the twenty-fifth annual ACM symposium on Parallelism in algorithms and architectures (pp. 241-250). ACM.

Deutsch Y, Golany B, Goldberg N, Rothblum U G (2013) Inspection games with local and global allocation bounds. *Naval Research Logistics (NRL).* 60(2):125-140.

Diamond H (1982) Minimax policies for unobservable inspections. *Mathematics of Operations Research.* 7(1):139-153.

Dresher M (1962) A sampling inspection problem in arms control agreements: A game-theoretic analysis. (No. RM-2972-ARPA). RAND CORP SANTA MONICA CA.

Dubey P, Mas-Colell A, Shubik M (1980) Efficiency properties of strategies market games: An axiomatic approach. *Journal of Economic Theory.* 22(2):339-362.

Ferguson T S, Melolidakis C (1998) On the inspection game. *Naval Research Logistics (NRL).* 45(3):327-334

Fokkink R, Lindelauf R (2013) The Application of Search Games to Counter Terrorism Studies. In Handbook of Computational Approaches to Counterterrorism (pp. 543-557). Springer New York.

Friedman D (1991) Evolutionary games in economics. *Econometrica: Journal of the Econometric Societ.* 637-666.

Friedman D (1998) On economic applications of evolutionary game theory. *Journal of Evolutionary Economics.* 8(1):15-43.

Garnaev A Y (1994) A remark on the customs and smuggler game. *Naval Research Logistics (NRL).* 41(2):287-293.

Gintis H (2000) *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Princeton university press.

Giovannoni F, Seidmann D J (2014) Corruption and power in democracies. *Social Choice and Welfare*. 42(3):707-34.

Gomes D A, Mohr J, Souza R R (2010) Discrete time, finite state space mean field games. *Journal de mathmatiques pures et appliques*. 93(3):308-328.

Gomes D A, Mohr J, Souza R R (2013) Continuous time finite state mean field games. *Applied Mathematics & Optimization*. 68(1):99-143.

Gomes D, Velho R M, Wolfram M T (2014) Socio-economic applications of finite state mean field games. *Phil. Trans. R. Soc. A*. 372(2028), 20130405.

Gomes D A, Saude J (2014) Mean field games modelsa brief survey. *Dynamic Games and Applications*. 4(2):110-154.

Green P (2004) The Sixteen Satires. Penguin UK.

Ho H M, Ouaknine J (2015, April) The cyclic-routing UAV problem is PSPACE-complete. In International Conference on Foundations of Software Science and Computation Structures (pp. 328-342). Springer Berlin Heidelberg.

Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics*. Cambridge University Press: Cambridge, UK.

Hofbauer J, Sigmund K (2003) Evolutionary game dynamics. *Bulletin of the American Mathematical Society*. 40(4):479-519.

Hochbaum D S, Lyu C, Ordez F (2014) Security routing games with multivehicle Chinese postman problem. *Networks*. 64(3):181-191.

Holte R, Mok A, Rosier L, Tulchinsky I, Varvel D (1989, January) The pinwheel: A real-time scheduling problem. In System Sciences, 1989. Vol. II: Software Track, Proceedings of the Twenty-Second Annual Hawaii International Conference on (Vol. 2, pp. 693-702). IEEE.

Hohzaki R (2007) An inspection game with multiple inspectees. *European Journal of Operational Research*. 178(3):894-906.

Hohzaki R (2013) Inspection games. *Wiley Encyclopedia of Operations Research and Management Science*.

Hurwicz L (2008) “But Who Will Guard the Guardians?”. *The American Economic Review*. 98(3):577-585

Huang M, Malham R P, Caines P E (2006) Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems.* 6(3):221-252.

Huang M (2010) Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM Journal on Control and Optimization.* 48(5):3318-3353.

Jain A K (2001) Corruption: A review. *Journal of economic surveys.* 15(1):71-121.

Jordan D W, Smith P (2007) *Nonlinear Ordinary Differential Equations: An Introduction for Scientists and Engineers.* Oxford University Press: Oxford, UK.

Jovanovic B, Rosenthal R W (1988) Anonymous sequential games. *Journal of Mathematical Economics.* 17(1):77-87.

Kamien M I, Schwartz N L (1991) *Dynamic optimisation. The calculus of variations and optimal control in economics and management.* North Holland, New York.

Katsikas S, Kolokoltsov V, Yang W (2016) Evolutionary Inspection and Corruption Games. *Games.* 7(4), 31.

Kolokoltsov V N (2010, July) Nonlinear Markov games. *In Proceedings of the 19th MTNS Symposium.*

Kolokoltsov V N (2010). *Nonlinear Markov processes and kinetic equations (Vol. 182).* Cambridge University Press.

Kolokoltsov V N, Malafeyev O A (2010) *Understanding Game Theory: Introduction to the Analysis of Many Agent Systems with Competition and Cooperation.* World Scientific, Singapore.

Kolokoltsov V N (2011) *Markov processes, semigroups, and generators (Vol. 38).* Walter de Gruyter.

Kolokoltsov V N, Yang W (2012) Turnpike theorems for Markov games. *Dynamic Games and Applications.* 2(3):294-312

Kolokoltsov V N (2012) Nonlinear Markov games on a finite state space (mean-field and binary interactions). *International Journal of Statistics and Probability.* 1(1).

Kolokoltsov V N (2014) The evolutionary game of pressure (or interference), resistance and collaboration. arXiv preprint arXiv:1412.1269. Available online: https://arxiv.org/abs/1412.1269 (accessed on 3 December 2014)

Kolokoltsov V N, Malafeyev O A (2015) Mean-field-game model of corruption. *Dynamic Games and Applications*. 1-14.

Kolokoltsov V N, Bensoussan A (2016) Mean-field-game model for Botnet defense in Cyber-security. *Applied Mathematics & Optimization*. 74(3):669-692.

Lambert-Mogiliansky A, Majumdar M, Radner R (2008) Petty corruption: A game-theoretic approach. *International Journal of Economic Theory*. 4(2):273-297.

Lasry J M, Lions P L (2007) Mean field games. *Japanese journal of mathematics*. 2(1):229-260.

Lee J H, Sigmund K, Dieckmann U, Iwasa Y (2015) Games of corruption: How to suppress illegal logging. *Journal of theoretical biology*. 367:1-3.

Lin K Y, Atkinson M P, Chung T H, Glazebrook K D (2013) A graph patrol problem with random attack times. *Operations Research*. 61(3):694-710.

Lin K Y, Atkinson M P, Glazebrook K D (2014) Optimal patrol to uncover threats in time when detection is imperfect. *Naval Research Logistics (NRL)*. 61(8):557-576.

Malafeyev O A, Redinskikh N D, Alferov G V (2014) Electric circuits analogies in economics modeling: Corruption networks. In Emission Electronics (ICEE), 2014 2nd International Conference on (pp. 1-5). IEEE.

Maschler M (1966) A price leadership method for solving the inspector's non-constant-sum game. *Naval Research Logistics (NRL)*. 13(1):11-33.

Maschler M (1967) The inspector's non-constant-sum game: Its dependence on a system of detectors. *Naval Research Logistics (NRL)*. 14(3):275-290

Morris P (2012) *Introduction to game theory*. Springer Science & Business Media.

Morse P M, Kimball G E (1951) *Methods of Operations Research*. MIT Press and Wiley, New York.

Myerson R B (2013) *Game theory*. Harvard university press.

Nikolaev P V (2014) Corruption suppression models: The role of inspectors moral level. *Computational Mathematics and Modeling*. 25(1):87-102.

Osborne M J, Rubinstein A (1994) *A course in game theory*. MIT press.

Papadaki K, Alpern S, Lidbetter T, Morton A (2016) Patrolling a border. *Operations Research*. 64(6):1256-1269.

Pita J, Jain M, Marecki J, Ordez F, Portway C, Tambe M, ..., Kraus S (2008, May) Deployed ARMOR protection: the application of a game theoretic model for security at the Los Angeles International Airport. In Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems: industrial track (pp. 125-132). International Foundation for Autonomous Agents and Multiagent Systems.

Ross S M (2014) *Introduction to stochastic dynamic programming.* Academic press.

Rothenstein D, Zamir S (2002) Imperfect inspection games over time. *Annals of Operations Research.* 109(1):175-192.

Rubinstein A (1986) Finite automata play the repeated prisoner's dilemma. *Journal of economic theory.* 39(1):83-96.

Sakaguchi M A (1977) A sequential allocation game for targets with varying values. *Journal of the Operations Research Society of Japan.* 20(3):182-193

Samuelson L (2002) Evolution and game theory. *The Journal of Economic Perspectives.* 16(2):47-66.

Sandler T, Arce D G (2007) Terrorism: a game-theoretic approach. *Handbook of defense economics.* 2:775-813.

Sandler T, Siqueira K (2009) Games and terrorism: Recent developments. *Simulation & Gaming.* 40(2):164-192.

Smith J M (1988) Evolution and the Theory of Games. In *Did Darwin Get It Right?* (pp. 202-215). Springer US.

Strogatz S H (2014) *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering.* Westview press.

Szabó G, Fath G (2007) Evolutionary games on graphs. *Physics reports.* 446(4):97-216.

Szechtman R, Kress M, Lin K, Cfir D (2008) Models of sensor operations for border surveillance. *Naval Research Logistics (NRL).* 55(1):27-41.

Taylor C, Fudenberg D, Sasaki A, Nowak M A (2004) Evolutionary game dynamics in finite populations. *Bulletin of mathematical biology.* 66(6):1621-1644.

Tembine H, Le Boudec J Y, El-Azouzi R, Altman E (2009, May) Mean field asymptotics of Markov decision evolutionary games and teams. In Game Theory for Networks, 2009. GameNets' 09. International Conference on (pp. 140-150). IEEE.

Thomas M U, Nisgav Y (1976) An infiltration game with time dependent payoff. *Naval Research Logistics (NRL)*. 23(2):297-302.

Von Neumann J, Morgenstern O (2007) *Theory of games and economic behavior*. Princeton university press.

Von Stengel B (1991) Recursive Inspection Games. Report No. S 9106; Computer Science Faculty; Armed Forces University: Munich, Germany, 1991.

Von Stengel B (2016) Recursive Inspection Games. *Mathematics of Operations Research*. 41(3):935-952

Washburn A R (1982) On patrolling a channel. *Naval Research Logistics Quarterly*. 29(4):609-615.

Washburn A (2010) Barrier games. *Military Operations Research*. 15(3):31-41.

Weibull J W (1997) *Evolutionary game theory*. MIT press.

Zaslavski A J (2006) *Turnpike properties in the calculus of variations and optimal control*. (VOL. 80). Springer Sceince & Business Media.

Zeeman E C (1980) Population dynamics from game theory. In Global Theory of Dynamical Systems (pp. 471-497). Springer Berlin, Heidelberg.

Zoroa N, Fernández-Sáez M J, Zoroa P (2012) Patrolling a perimeter. *European Journal of Operational Research*. 222(3):571-582.