Check for updates

Where Do Mistakes Lead? A Survey of Games with Incompetent Players

Thomas Graham[1] · Maria Kleshnina[2,3] · Jerzy A. Filar[1]

© The Author(s) 2022

Abstract

Mathematical models often aim to describe a complicated mechanism in a cohesive and simple manner. However, reaching perfect balance between being simple enough or overly simplistic is a challenging task. Frequently, game-theoretic models have an underlying assumption that players, whenever they choose to execute a specific action, do so perfectly. In fact, it is rare that action execution perfectly coincides with intentions of individuals, giving rise to behavioural mistakes. The concept of incompetence of players was suggested to address this issue in game-theoretic settings. Under the assumption of incompetence, players have non-zero probabilities of executing a different strategy from the one they chose, leading to stochastic outcomes of the interactions. In this article, we survey results related to the concept of incompetence in classic as well as evolutionary game theory and provide several new results. We also suggest future extensions of the model and argue why it is important to take into account behavioural mistakes when analysing interactions among players in both economic and biological settings.

Keywords Incompetence · Execution errors · Learning · Game theory · Stochastic games · Evolutionary games

✉ Jerzy A. Filar
   j.filar@uq.edu.au

   Thomas Graham
   t.graham2@uqconnect.edu.au

   Maria Kleshnina
   maria.kleshnina@iast.fr

[1] School of Mathematics and Physics, University of Queensland, Brisbane, Australia

[2] Institute of Science and Technology Austria (IST Austria), Klosterneuburg, Austria

[3] Institute for Advanced Studies in Toulouse, Toulouse, France

🕲 Birkhäuser

1 Introduction

In classical non-cooperative games, the payoffs (or outcomes) are determined directly by the players' choice of strategies. In reality, however, a player may not be capable of executing his or her chosen strategy due to a lack of skill that we shall refer to as *incompetence*.

In this paper, we survey a relatively recent line of research that is predicated on the assumption that player incompetence is a real and ubiquitous phenomenon that deserves deeper investigation. In the process, we also present a few recent results which, to the best of our knowledge, have not been reported elsewhere. Since we regard the topic of incompetent games as still being in its infancy, the main objective of this paper is to stimulate further research on this subject.

Naturally, players' incompetence inherently complicates a game. To prevent the added complexity from becoming unmanageable, we must impose some assumptions on the information domain and the structural form via which incompetence manifests itself. In the development so far, the following key conceptual assumption has been imposed:

> [A1] *Incompetence manifests itself as a set of probability distributions on sets of actions available to one or more players.*

While [A1] is restrictive in some ways, it allows us to recover a classical competent game as a special case of an incompetent game. Namely, a game where all of the above incompetence distributions are degenerate and execute the selected actions with certainty.

This immediately raises the possibility of parametrising the *level of competence* or, equivalently, *level of skill* of a player by the 'proximity' to such a, fully competent, degenerate distribution. It also opens up the possibility of players 'learning' by reducing their levels of incompetence (equivalently, increasing their skill). We note that this kind of learning is essentially different from both the discovery statistical learning and the imitation machine learning.

To date, the topic of incompetent games has evolved along two distinct, but conceptually related, directions. The first of these is the study of classical non-cooperative games under the assumption that at least one player is incompetent. The second is the study of evolutionary incompetent games.

In the case of classical games under incompetence, existing analyses assumed that all probability distribution capturing incompetence are mutually known by all players. This is plausible in the case of players who are familiar with others' past performance (e.g. tennis players on the international tour circuit). However, there is certainly scope for relaxing that assumption. Such relaxations may give rise to interesting repeated versions of these games and the natural trade-off between the so-called problem of "exploration versus exploitation".

The preceding "mutually known" assumption is not explicitly needed in the evolutionary incompetent games. It is also conceptually challenging to ascribe consciousness of such distributions to individual animals or bacteria. Nonetheless, within the evolutionary paradigm, it is reasonable to assume that the emerging equilibrium frequencies of species types have incorporated the mutual incompetence uncertainties in their adaptation to the ecosystem. The uncertainties stemming from the incompetence are thus simply built into the replicator dynamics of the game.

This review paper is structured as follows. We introduce incompetence first in classical non-cooperative games and then in evolutionary games. In Sect. 2, specifically in Sects. 2.1-2.4, we provide an overview of the formal definitions and results on games with incompetent players in classical settings. Then, in Sect. 2.5 we introduce a new concept of incremental

learning in the games with incompetent players and derive some of its properties. In Sect. 3, we define evolutionary games with incompetence and provide an overview of results on these games. Additionally, we provide a rationale for the importance of considering these games in biological settings. We conclude by suggesting possible directions for future extensions of games with incompetence.

1.1 Incompetent Classical Non-cooperative Games

Chronologically, Beck and Filar [10] introduced incompetence to matrix games and Beck *et* al. [9] introduced incompetence to bimatrix games. Essentially, by quantifying a player's tendency to accidentally deviate from their selected actions, these authors represent incompetence as stochastic matrices that can be used to account for these deviations. The application of incompetent classical games to military planning is discussed in [8] and [7].

We note that the notion of incompetence introduced here is superficially similar to several concepts used to measure the sensitivity of equilibria to changes in a game's parameters. For example, Selten [63] imagines players having "trembling hands" that cause them to accidentally execute unintended actions with negligible probability. This is used to refine the equilibrium solution concept in extensive-form games by defining trembling hand perfect equilibria. However, unlike the notion of incompetence, a trembling hand is not intended to model players making mistakes with arbitrary or even prescribed probabilities (e.g. a tennis player who routinely places a 'passing shot' in the opponent's hitting zone).

While the concepts and some of the results are generalisable to N-person non-cooperative games, the setting of classic two-player games—meaning matrix and bimatrix games—is a natural starting point for the introduction of incompetence to game-theoretic models. Larkey *et* al. [45], when discussing the game "Sum Poker", observe that there are several cognitive and physical limitations that might prevent a player from finding or implementing optimal strategies. Then, seeking to classify the specific difficulties a player encounters, they propose a typology of skills consisting of:

- *Strategic Skill*, the ability to select which games should be played,
- *Planning Skill*, the ability to develop a desirable strategy within a game, and
- *Execution Skill*, the ability to execute desired actions throughout a game.

Although Larkey *et* al. [45] apply this typology to experimentally compare different strategies in "Sum Poker" under different skill limitations, a precise mathematical formulation of skill is not provided. The notion of incompetence mainly addresses the issue of execution skill as it quantifies accidental (ipso facto, unintended) deviations from a player's chosen strategy.

The concept of incompetence is a useful modelling tool in traditional game-theoretic settings because it captures a player's inability to precisely control the outcome of their actions. Necessarily, real-world strategic interactions are often complicated and a player's intentions might not be perfectly realised due to noise from their environment. For instance, a tennis player is unable to control the exact trajectory of a shot and might sometimes make consequential mistakes. Similarly, in the economic context, it is conceivable that a firm in the classic Cournot oligopoly model is unable to guarantee the quantity of goods produced, perhaps due to sporadic errors occurring during production. The latter may reflect flaws in the firm's quality control regime which in itself is related to its level of competence[1]. The players in these situations must accept some degree of variability in the outcomes of their actions and incompetence is a method capable of accounting for this variability.

---

[1] Arguably, the famous "Toyota Production System (TPS)" could be seen as that firm's highly successful effort to reduce its level of incompetence.

In this paper, in the context of classical games, we will review the introduction of incompetence and present several related, unpublished, results. Moreover, we will discuss and solve a simple model for incrementally learning to decrease incompetence during a repeatedly-played incompetent matrix game.

1.2 Incompetent Evolutionary Games

Game theory applied to populations of species has branched out from non-cooperative game theory and evolved into an independent field called *evolutionary game theory* [28,56,68]. The setup of evolutionary games differs from the classical games in the very basic assumption rationality of players. Obviously, one cannot expect rational behaviour from individual bacteria or lions. However, the selection strength still acts rationally, which follows from the classic prediction that the fittest survives [69]. Hence, in evolutionary games players, or animals, do not make conscious choice of strategies to play as humans in economic settings. Instead, individuals are born with a strategy predefined by their parents. The competition then happens at a genetic level, which encodes the strategy. Incidentally, the question of how cooperation evolves in biology is still open and, possibly, related to the coexistence of different strategies at equilibrium. We note that in a recent paper [2] thermodynamics was invoked to shed light on cooperation in evolutionary games.

The concept of incompetence was considered in biological settings when studying the evolution of social behaviour [36–40]. That is, incompetence now acts at the selection level rather than organisms themselves. Then, incompetence can be seen as behavioural plasticity leading to mixed strategies executed at a genetic level. As a result, we do not require organisms to be aware of probability distributions of all genes as well as their mistakes. Instead, selection forces act upon these distributions driving the competition among types. Such plasticity can be of variable degree, depending on the environmental conditions and adaptations of organisms. The latter will correspond to the level of incompetence discussed above.

Naturally, the idea of behavioural plasticity or stochasticity is not novel in the field of evolutionary games and recently became one of the foci in the field. There are many approaches considered in biological settings such as genetic mutations [12,57,70,72], learning processes [23,29,41,42,51,57,64], adaptation dynamics [47], phenotypic plasticity [15], noise in continuous and discrete-time replicator dynamics [4,6,19,22] and environmental fluctuations [75,81]. Thus, the notion of incompetence of players merely fills a new niche where behavioural stochasticity is only induced at the moment of interactions.

Let us demonstrate this concept on a well-studied example of a Hawk-Dove game [67]. Imagine that individuals in a well-mixed large population compete for some resource. Two behavioural strategies are available in a population: a Hawk (aggressive) strategy and a Dove (passive) strategy. That is, Hawks fight for the resource, while Doves prefer to share equally and flee when attacked. This game has a payoff matrix of the same structure as a Chicken or Snowdrift games. As in these classical examples, in an equilibrium, both strategies stably coexist. The game can be illustrated as an interaction between a naturally aggressive person and a naturally passive one. Of course, a counter-attack in response to aggression is not something one naturally expects from a passive person. However, behavioural plasticity induced by incompetence can lead to situations when a passive player responses aggressively or an aggressive player flees instead of fighting. While their strategies did not change as such, the behaviour exhibited by individuals was altered. In [37], it was shown that such an assumption may lead to different evolutionary outcomes and may change the way selection

is realised. In this survey, we discuss these results and demonstrate them on an example of a biological game.

2 Incompetence in Classical Non-cooperative Games

2.1 Games without Incompetence

An $m \times n$ *bimatrix game* Γ consists of a pair of *action sets* $\mathcal{A} = \{1, 2, \ldots, m\}$ and $\mathcal{B} = \{1, 2, \ldots, n\}$ and a pair of *reward matrices* $R^1 = (r^1(i, j)) \in \mathbb{R}^{m \times n}$ and $R^2 = (r^2(i, j)) \in \mathbb{R}^{m \times n}$. Here, throughout this section only, we use non-standard notation for matrix entries to accommodate additional symbols associated with incompetent games. After an action $i \in \mathcal{A}$ is selected by Player 1 and an action $j \in \mathcal{B}$ is selected by Player 2, they receive rewards according to the matrix entries $r^1(i, j)$ and $r^2(i, j)$, respectively. A *(mixed) strategy* extends this behaviour to allow for randomised action selection. Specifically, Player 1 chooses a strategy from the probability simplex

$$\mathbf{X} := \left\{ \mathbf{x} = (x_i)_{i=1}^m \in \mathbb{R}^m : \sum_{i=1}^m x_i = 1 \text{ and } x_i \geq 0, \forall i = 1, 2, \ldots, m \right\} \tag{1}$$

over $\mathcal{A}$ and Player 2 chooses a strategy from the probability simplex

$$\mathbf{Y} := \left\{ \mathbf{y} = (y_j)_{j=1}^n \in \mathbb{R}^n : \sum_{j=1}^n y_j = 1 \text{ and } y_j \geq 0, \forall j = 1, 2, \ldots, n \right\} \tag{2}$$

over $\mathcal{B}$. The resulting strategy profile $(\mathbf{x}, \mathbf{y}) \in \mathbf{X} \times \mathbf{Y}$ yields an expected reward of

$$v^k(\mathbf{x}, \mathbf{y}) := \mathbf{x} R^k \mathbf{y}^T \tag{3}$$

to Player $k \in \{1, 2\}$ where $\mathbf{y}^T$ is the transpose of $\mathbf{y}$. We want to find the *(Nash) equilibria* of Γ, which capture the notion of a stable strategy profile. Precisely, $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{X} \times \mathbf{Y}$ is an equilibrium whenever it is resilient to unilateral deviations or, equivalently,

$$\mathbf{x} R^1 (\mathbf{y}^*)^T \leq \mathbf{x}^* R^1 (\mathbf{y}^*)^T \quad \text{and} \quad \mathbf{x}^* R^2 \mathbf{y}^T \leq \mathbf{x}^* R^2 (\mathbf{y}^*)^T \tag{4}$$

for all $\mathbf{x} \in \mathbf{X}$ and $\mathbf{y} \in \mathbf{Y}$. Nash [54], in a seminal contribution to game theory, proves that every game with finitely-many players and actions has an equilibrium point. Vorob'ev [79] shows that, in bimatrix games, the set of equilibria is the union of a finite collection of convex sets. Specifically, these are called maximal Nash subsets and are the largest equilibrium-containing sets that are closed under interchanging a player's strategies. The extreme points of a maximal Nash subset are called extreme equilibria and are associated with paired non-singular square submatrices of R^1 and R^2 called kernels (see Kuhn [43]).

A bimatrix game satisfying the *zero-sum* property: $r^1(i, j) = -r^2(i, j)$ for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, is called a *matrix game* and is described by the single matrix $R := R^1 = -R^2$. Note that, when Γ is a matrix game, every equilibrium $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{X} \times \mathbf{Y}$ achieves the same reward $\mathrm{val}(\Gamma) := v^1(\mathbf{x}^*, \mathbf{y}^*) = -v^2(\mathbf{x}^*, \mathbf{y}^*)$, which is called the *value of the game* Γ. Moreover, in recognition of von Neumann's [55] celebrated minimax theorem showing

$$\mathrm{val}(\Gamma) = \max_{\mathbf{x} \in \mathbf{X}} \min_{\mathbf{y} \in \mathbf{Y}} \mathbf{x} R \mathbf{y}^T = \min_{\mathbf{y} \in \mathbf{Y}} \max_{\mathbf{x} \in \mathbf{X}} \mathbf{x} R \mathbf{y}^T, \tag{5}$$

the component strategies $\mathbf{x}^*$ and $\mathbf{y}^*$ of an equilibrium are often called *(minimax) optimal strategies*.

A *completely mixed equilibrium* $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{X} \times \mathbf{Y}$ of a bimatrix game Γ is an equilibrium under which every action is played with non-zero probability. If Γ has only completely mixed equilibria, then it is called a *completely mixed game* and has only a single (completely mixed) equilibrium (see Kaplansky [35] and Raghavan [62]). Additionally, if Γ has a maximal Nash subset containing only completely mixed equilibria, then it is called a *weakly completely mixed game*. Jurg *e*t al. [34] prove that a weakly completely mixed game contains a unique completely mixed equilibrium.

2.2 Games with Incompetence

Beck *e*t al. [9] introduce incompetence to a bimatrix game Γ by allowing players to accidentally deviate from their intended actions. Specifically, after a pair of actions is selected from $\mathcal{A} \times \mathcal{B}$, incompetence randomly determines an executed action profile—also from $\mathcal{A} \times \mathcal{B}$—according to a predefined probability distribution. This distribution is represented by the *incompetence matrices* $Q^1 := (q^1(i, i')) \in \mathbb{R}^{m \times m}$ and $Q^2 := (q^2(j, j')) \in \mathbb{R}^{n \times n}$. After Player 1 and Player 2 select the actions $i \in \mathcal{A}$ and $j \in \mathcal{B}$, they execute the actions $i' \in \mathcal{A}$ and $j' \in \mathcal{B}$ with probability $q^1(i, i')$ and $q^2(j, j')$, respectively. The notation $\Gamma_{Q^1 Q^2}$ denotes the game Γ played under incompetence. If the incompetence matrices Q^1 and Q^2 are unambiguous, we will often replace the subscript "$Q^1 Q^2$" with simply "Q" (e.g. Γ_Q instead of $\Gamma_{Q^1 Q^2}$).

Note that the original incompetence framework described by Beck *e*t al. [9] allows the players' sets of selectable and executable actions to differ. This is especially useful for modelling actions that can be executed with variable quality. Beck and Filar [10] give an example of a capability acquisition game in which a defender, after selecting the action "Conventional Defence", may execute either "Good Conventional Defence" or "Bad Conventional Defence". Here, for the sake of notational simplicity, we assume that a player's sets of selectable and executable actions coincide.

Suppose that Player 1 selects the action $i \in \mathcal{A}$ and Player 2 selects the action $j \in \mathcal{B}$. The expected reward received by Player $k \in \{1, 2\}$ under incompetence is

$$r_Q^k(i, j) := \sum_{i'=1}^{m} \sum_{j'=1}^{n} q^1(i, i') r^k(i', j') q^2(j, j'). \tag{6}$$

Hence, Γ_Q can be treated as another bimatrix game with the incompetent reward matrix $R_Q^k := (r_Q^k(i, j)) \in \mathbb{R}^{m \times n}$ belonging to Player $k \in \{1, 2\}$. Clearly, we have

$$R_Q^k = Q^1 R^k (Q^2)^T. \tag{7}$$

Note that, as an immediate consequence of (7), an incompetent game derived from a matrix game is also a matrix game. The expected reward granted to Player $k \in \{1, 2\}$ in Γ_Q is

$$v_Q^k(\mathbf{x}, \mathbf{y}) := \mathbf{x} R_Q^k \mathbf{y}^T = \mathbf{x} Q^1 R_Q^k (Q^2)^T \mathbf{y}^T \tag{8}$$

for each $(\mathbf{x}, \mathbf{y}) \in \mathbf{X} \times \mathbf{Y}$.

Beck *e*t al. [9] are not only interested in games with static incompetence, but also dynamic games wherein players are able to vary their incompetence. This is captured by a pair of *learning trajectories* $Q^1 : [0, 1] \to \mathbb{R}^{m \times m}$ and $Q^2 : [0, 1] \to \mathbb{R}^{n \times n}$. Then, for each pair of *learning parameters* $\lambda, \mu \in [0, 1]$, the corresponding incompetent game $\Gamma_Q(\lambda, \mu)$ has Player 1's incompetence matrix defined as $Q^1(\lambda)$ and Player 2's incompetence matrix defined as $Q^2(\mu)$. Equivalently, $\Gamma_Q(\lambda, \mu)$ is the bimatrix game with reward matrices

$$R_Q^k(\lambda, \mu) := R_{Q^1(\lambda)Q^2(\mu)}^k = Q^1(\lambda)R^k Q^2(\mu)^T \tag{9}$$

for each Player $k \in \{1, 2\}$. Although the family of parameterised incompetent games $\Gamma_Q(\lambda, \mu)$ is interesting in its own right (see Sect. 2.4), it is also an essential building-block used to construct dynamic learning games (see Sect. 2.5).

Example (attack-defence game with incompetence) Consider, as an example, a matrix game Γ played between two aeroplane pilots—labelled "Attacker" (Player 1) and "Defender" (Player 2)—competing over three sites. The attacker wants to destroy a site and the defender wants to prevent this from occurring. Precisely, we have the action sets $\mathcal{A} = \{1, 2, 3\}$ and $\mathcal{B} = \{1, 2, 3\}$ where, for each $i, j \in \{1, 2, 3\}$, Player 1's action i means "Attack Site i" and Player 2's action j means "Defend Site j". A successful attack occurs if and only if the defending pilot does not anticipate the attacking pilot's destination or, equivalently, the executed action profile $(i, j) \in \mathcal{A} \times \mathcal{B}$ has $i \neq j$. The attacker receives a reward $\nu_1 = 3$ when Site 1 is destroyed, $\nu_2 = 4$ when Site 2 is destroyed, and $\nu_3 = 5$ when Site 3 is destroyed. The corresponding utility matrix is

$$R = \begin{pmatrix} 0 & \nu_1 & \nu_1 \\ \nu_2 & 0 & \nu_2 \\ \nu_3 & \nu_3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 3 \\ 4 & 0 & 4 \\ 5 & 5 & 0 \end{pmatrix} \tag{10}$$

where $R = R^1 = -R^2$. The game value of Γ is $120/47 \approx 2.55$ and its (unique) equilibrium has the attacking strategy $\mathbf{x}^* = 1/47(\, 20\ 15\ 12\,)$ and the defending strategy $\mathbf{y}^* = 1/47(\, 7\ 17\ 23\,)$.

We use incompetence to capture the pilots' navigation skills and their propensity to arrive at an incorrect site after getting lost. Define their learning trajectories $Q^1, Q^2 : [0, 1] \to \mathbb{R}^{3\times3}$ where, for each $\lambda, \mu \in [0, 1]$, we set

$$Q^1(\lambda) = \frac{1}{3}J_3(1 - \lambda) + I_3\lambda \quad \text{and} \quad Q^2(\mu) = \frac{1}{3}J_3(1 - \mu) + I_3\mu \tag{11}$$

where $J_n \in \mathbb{R}^{n\times n}$ is the $n \times n$ all-ones matrix and $I_n \in \mathbb{R}^{n\times n}$ is the $n \times n$ identity matrix. Note that $Q^k = 1/n J_n$ is called uniform incompetence and $Q^k = I_n$ is called complete competence
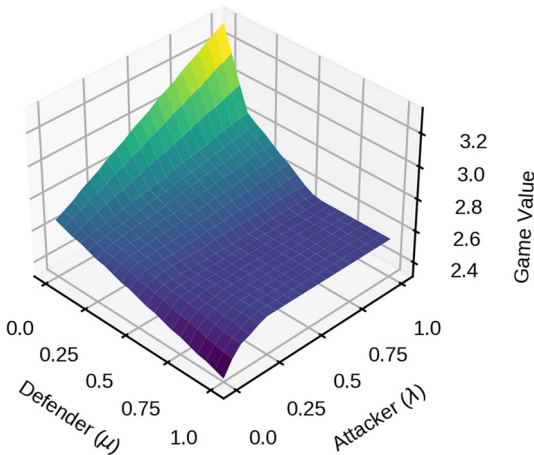

Fig. 1 The game value of the incompetent attack-defence game $\Gamma_Q(\lambda, \mu)$ for each pair of learning parameters $\lambda, \mu \in [0, 1]$

[9]. The resulting incompetent reward matrices at $\lambda = \mu = 0$ and $\lambda = \mu = 1/2$ are

$$R_Q(0, 0) = \frac{8}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad R_Q(1/2, 1/2) = \frac{1}{12}\begin{pmatrix} 23 & 31 & 30 \\ 37 & 24 & 35 \\ 42 & 41 & 25 \end{pmatrix}, \quad (12)$$

respectively. The game value of $\Gamma_Q(0, 0)$ is $8/3$ and the game value of $\Gamma_Q(1/2, 1/2)$ is $835/324 \approx 2.58$. Moreover, since complete competence is achieved at $\lambda = \mu = 1$ and $\Gamma_Q(1, 1) = \Gamma$, we already know that the game value of $\Gamma_Q(1, 1)$ is $120/47 \approx 2.55$. Thus, it appears that the parameterised incompetent games $\Gamma_Q(\lambda, \mu)$ move in the defender's favour as the learning parameters are increased along $\lambda = \mu$.

A clearer picture of the game value's dependence on these learning parameters is achieved in Fig. 1 by plotting the function $(\lambda, \mu) \mapsto \text{val}(\Gamma_Q(\lambda, \mu))$ on the domain $[0, 1] \times [0, 1]$. Note that, in this specific example, $(\lambda, \mu) \mapsto \text{val}(\Gamma_Q(\lambda, \mu))$ is piecewise linear and non-decreasing (or non-increasing) in the variable λ (or μ). This means that learning is beneficial for the attacking and defending player when their opponent's incompetence remains fixed. Furthermore, the game value plateaus on the region $[11/47, 1] \times [26/47, 1]$ indicating that the attacker reaches their "maximum useful skill' at $\lambda = 11/47 \approx 0.23$ and the defender reaches their "maximum useful skill" at $\mu = 26/47 \approx 0.55$. Interestingly, $\Gamma_Q(\lambda, \mu)$ is also completely mixed on $(11/47, 1] \times (26/47, 1]$, which suggests a connection between complete mixedness and this game value plateau. We will further explore the general properties of parameterised incompetent games in Sect. 2.4.

2.3 Executable Strategies

Although [9] and [10] view incompetence as modifying a player's reward matrix, it is also possible to view it as modifying their strategy spaces. Here, we return to the setting of a static incompetent game Γ_Q with incompetence matrices $Q^1 \in \mathbb{R}^{m \times m}$ and $Q^2 \in \mathbb{R}^{n \times n}$. Note that, after Player 1 selects a strategy $\mathbf{x} \in \mathbf{X}$ (or Player 2 selects a strategy $\mathbf{y} \in \mathbf{Y}$), the resulting executed strategy is $\mathbf{x}Q^1$ (or $\mathbf{y}Q^2$) after incompetence is included. What strategies are the players able to execute? Well, Player 1 and Player 2 are able to execute the strategies in

$$\mathbf{E}^1(Q^1) := \{\mathbf{x}Q^1 : \mathbf{x} \in \mathbf{X}\} \quad \text{and} \quad \mathbf{E}^2(Q^2) := \{\mathbf{y}Q^2 : \mathbf{y} \in \mathbf{Y}\}, \quad (13)$$

respectively. We call $\mathbf{E}^k(Q^k)$ the *executable strategy space* belonging to Player $k \in \{1, 2\}$ and, when the incompetence matrices are unambiguous, we simply write $\mathbf{E}^k$ instead. Importantly, from the perspective of an outside observer who only sees that players have executed strategies from $\mathbf{E}^1$ and $\mathbf{E}^2$, we would be unable to distinguish whether they are playing the competent game Γ or the incompetent game Γ_Q. Figure 2 shows some of the executable strategy spaces within the previously discussed attack-defence game with incompetence. Note that the transition between executable strategy spaces can be more complicated than the "growing" seen in Fig. 2.

What is the connection between equilibria of Γ_Q in $\mathbf{X} \times \mathbf{Y}$ and equilibria of Γ in $\mathbf{E}^1 \times \mathbf{E}^2$? Theorem 1 gives conditions under which an equilibrium in the competent game Γ can be converted into an equilibrium in the incompetent game Γ_Q, and vice versa. Theorem 1(i) implies that an equilibrium of Γ in $\mathbf{E}^1 \times \mathbf{E}^2$ is always executed by an equilibrium of Γ_Q. Meanwhile, Theorem 1(ii) implies that there exists an equilibrium of Γ in the interior of $\mathbf{E}^1 \times \mathbf{E}^2$ provided that Γ_Q is weakly completely mixed.

Lemma 1 *If Γ_Q is a weakly completely mixed incompetent bimatrix game, then its incompetence matrices Q^1 and Q^2 are non-singular.*
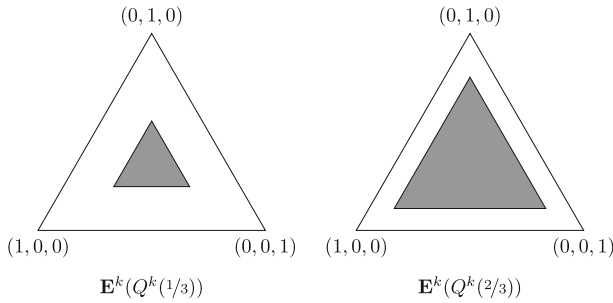
Fig. 2 The executable strategy spaces belonging to Player $k \in \{1, 2\}$ in the previously introduced attack-defence game at $\lambda = \mu = 1/3$ and $\lambda = \mu = 2/3$

Proof Let $\mathbf{q}^1(i)$ denote the ith row of Q^1. Assuming that Q^1 is singular, there exists $I \subseteq \{1, 2, \ldots, m\}$ such that

$$\sum_{i \in I} \theta_i \mathbf{q}^1(i) = 0$$

for some non-zero coefficients $\theta_i \in \mathbb{R}$ with $i \in I$. Then, after right-multiplying by the all-ones row vector $\mathbf{1}_m^T \in \mathbb{R}^m$, we have $\sum_{i \in I} \theta_i = 0$. Certainly, since the entries of Q^1 are non-negative, we can partition the index the set I into non-empty subsets $I^+ = \{i \in I : \theta_i > 0\}$ and $I^- = \{i \in I : \theta_i < 0\}$ with $\sum_{i \in I^+} \theta_i = -\sum_{i \in I^-} \theta_i$.

Let $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{X} \times \mathbf{Y}$ be an equilibrium of Γ_Q. Define a constant $\alpha = -\min_{i \in I^-}\{x_i^*/\theta_i\}$ such that $x_i^* + \alpha\theta_i \geq 0$ for all $i \in I$ and $x_j^* + \alpha\theta_j = 0$ for some $j \in I$. We construct an alternative strategy $\mathbf{x}^\dagger \in \mathbf{X}$, which is also completely mixed, where

$$\mathbf{x}_i^\dagger = \begin{cases} x_i^* + \frac{1}{2}\alpha\theta_i, & i \in I, \\ x_i^*, & i \notin I, \end{cases}$$

for each $i = 1, 2, \ldots, m$. Observe that

$$\mathbf{x}^\dagger Q^1 = \sum_{i=1}^m x_i^\dagger \mathbf{q}^1(i) = \sum_{i \notin I} x_i^* \mathbf{q}^1(i) + \sum_{i \in I} \left(x_i^* + \frac{\alpha}{2}\theta_i\right)\mathbf{q}^1(i)$$

$$= \sum_{i=1}^m x_i^* \mathbf{q}^1(i) + \frac{\alpha}{2} \sum_{i \in I} \theta_i \mathbf{q}^1(i) = \mathbf{x}^* Q^1,$$

so $\mathbf{x}^*$ and $\mathbf{x}^\dagger$ result in identical expected rewards to Player 1 and Player 2 in Γ_Q and $(\mathbf{x}^\dagger, \mathbf{y}^*)$ is an equilibrium of Γ_Q. But, given that Γ_Q contains two distinct completely mixed equilibria $(\mathbf{x}^*, \mathbf{y}^*)$ and $(\mathbf{x}^\dagger, \mathbf{y}^*)$, it cannot be a weakly completely mixed game. After using a similar argument for the other incompetence matrix Q^2, the desired result follows by contraposition. □

Theorem 1 *Fix a strategy profile $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{X} \times \mathbf{Y}$ in an incompetent bimatrix game Γ_Q. Then,*

(i) *$(\mathbf{x}^*, \mathbf{y}^*)$ is an equilibrium in Γ_Q whenever $(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ is an equilibrium in Γ, and*
(ii) *$(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ is a (completely mixed) equilibrium in Γ whenever Γ_Q is weakly completely mixed and $(\mathbf{x}^*, \mathbf{y}^*)$ is a completely mixed equilibrium in Γ_Q.*

Proof (i) Assume that Player 1 possesses a profitable deviation from $(\mathbf{x}^*, \mathbf{y}^*)$ in Γ_Q such that $v_Q^1(\mathbf{x}, \mathbf{y}^*) > v_Q^1(\mathbf{x}^*, \mathbf{y}^*)$ for some $\mathbf{x} \in \mathbf{X}$. Then, observe that

$$v^1(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2) = \mathbf{x}^* Q^1 R^1 (Q^2)^T (\mathbf{y}^*)^T = \mathbf{x}^* R_Q^1 (\mathbf{y}^*)^T$$
$$< \mathbf{x} R_Q^1 (\mathbf{y}^*)^T = \mathbf{x} Q^1 R^1 Q^2 (\mathbf{y}^*)^T = v^1(\mathbf{x} Q^1, \mathbf{y}^* Q^2)$$

which contradicts the equilibrium inequalities for $(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ in Γ. After repeating a similar argument for Player 2, we obtain

$$v_Q^1(\mathbf{x}, \mathbf{y}^*) \le v_Q^1(\mathbf{x}^*, \mathbf{y}^*) \quad \text{and} \quad v_Q^2(\mathbf{x}^*, \mathbf{y}) \le v_Q^2(\mathbf{x}^*, \mathbf{y}^*),$$

meaning that $(\mathbf{x}^*, \mathbf{y}^*)$ is an equilibrium of Γ_Q.

(ii) We know that Player 1's strategy $\mathbf{x}^*$ makes Player 2 indifferent between their actions in Γ_Q, hence

$$\mathbf{x}^* R_Q^2 = \mathbf{x}^* Q^1 R^2 (Q^2)^T = v^2(\mathbf{x}^*, \mathbf{y}^*) \mathbf{1}_n^T,$$

where $\mathbf{1}_n \in \mathbb{R}^{1 \times n}$ is an all-ones row vector. Clearly, this is solved when $\mathbf{x}^* Q^1 R^2 = v^2(\mathbf{x}^*, \mathbf{y}^*) \mathbf{1}_n^T$ and this solution is unique as Q^2 is non-singular (by Lemma 1). This shows that $\mathbf{x}^* Q^1$ makes Player 2 indifferent between their actions in Γ and, by a similar argument, $\mathbf{y}^* Q^2$ makes Player 1 indifferent between their actions in Γ. Note that $\mathbf{x}^* Q^1$ and $\mathbf{y}^* Q^2$ are both completely mixed because the entries in $\mathbf{x}^*$ and $\mathbf{y}^*$ are strictly positive and (by non-singularity) the columns of Q^1 and Q^2 cannot contain only zeros. Thus, appealing to the indifference principle, we conclude that $(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ is an equilibrium in Γ. □

Corollary 1, which states that a completely mixed matrix game achieves the same game value as its competent counterpart, was originally presented by Beck and Filar [10]. Although they give a utility-centred argument based on Shapley and Snow's [66] game value formula, we give an alternative strategy-centred argument based on Theorem 1(ii). Note that, a generalisation of this result to bimatrix games is presented in [9]; however, we still choose to highlight the matrix game version for later discussion.

Corollary 1 *[10] If Γ_Q is a completely mixed incompetent matrix game, then Γ is a matrix game and $\mathrm{val}(\Gamma) = \mathrm{val}(\Gamma_Q)$.*

Proof Certainly, Γ is also a matrix game because $R^1 \ne -R^2$ directly implies $R_Q^1 \ne -R_Q^2$. If $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{X} \times \mathbf{Y}$ is the unique equilibrium of Γ_Q, then Theorem 1(ii) states that $(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ is an equilibrium of Γ. So,

$$\mathrm{val}(\Gamma) = v(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2) = \mathbf{x}^* Q^1 R (Q^2)^T (\mathbf{y}^*)^T$$
$$= \mathbf{x}^* R_Q (\mathbf{y}^*)^T = v_Q(\mathbf{x}^*, \mathbf{y}^*) = \mathrm{val}(\Gamma_Q),$$

as desired. □

Lastly, the result in Theorem 1(ii) can be extended to incompetent bimatrix games that are "almost" weakly completely mixed. Theorem 2 shows that, if Γ_Q can be approximated by a sequence of weakly completely mixed incompetent games, then Γ has an equilibrium in $\mathbf{E}^1 \times \mathbf{E}^2$.

Lemma 2 *If Γ_Q is a weakly completely mixed incompetent bimatrix game, then Γ is also weakly completely mixed.*

Proof We know from Theorem 1(ii) that there exists a completely mixed equilibrium $(\mathbf{x}^*, \mathbf{y}^*) \in \mathbf{E}^1 \times \mathbf{E}^2$ of Γ, so it lies in the interior of $\mathbf{E}^1 \times \mathbf{E}^2$. If Γ is not weakly completely mixed, then there exists another equilibrium $(\mathbf{x}^\dagger, \mathbf{y}^\dagger) \in \mathbf{X} \times \mathbf{Y}$ such that $(\mathbf{x}^*, \mathbf{y}^*)$ and $(\mathbf{x}^\dagger, \mathbf{y}^\dagger)$ belong to the same maximal Nash subset. Define the convex combination $(\mathbf{x}^\alpha, \mathbf{y}^\alpha)$ of these strategy profiles by

$$\mathbf{x}^\alpha = \alpha \mathbf{x}^* + (1 - \alpha)\mathbf{x}^\dagger \quad \text{and} \quad \mathbf{y}^\alpha = \alpha \mathbf{y}^* + (1 - \alpha)\mathbf{y}^\dagger$$

for each $\alpha \in [0, 1]$. Note that, because Nash subsets are closed under convex combinations, $(\mathbf{x}^\alpha, \mathbf{y}^\alpha)$ is an equilibrium of Γ for every $\alpha \in [0, 1]$. Moreover, for some $\alpha^* \in (0, 1]$, the strategy profile $(\mathbf{x}^\alpha, \mathbf{y}^\alpha)$ lies in the interior of $\mathbf{E}^1 \times \mathbf{E}^2$ for every $\alpha \in [0, \alpha^*)$. But, by Theorem 1(i) and Lemma 1, this means that $(\mathbf{x}^\alpha (Q^1)^{-1}, \mathbf{y}^\alpha (Q^2)^{-1})$ is a (completely mixed) equilibrium of Γ_Q for each $\alpha \in [0, \alpha^*)$. Given that this contradicts the uniqueness of a completely mixed equilibrium in the weakly completely mixed game Γ_Q, we conclude that Γ must also be weakly completely mixed. $\qquad\square$

Theorem 2 *Let $\{Q_\ell^1\}_{\ell=1}^\infty$ and $\{Q_\ell^2\}_{\ell=1}^\infty$ be sequences of incompetence matrices that converge to Q^1 and Q^2, respectively. Moreover, assume that $\Gamma_{Q_\ell} = \Gamma_{Q_\ell^1 Q_\ell^2}$ is weakly completely mixed for every $\ell = 1, 2, \ldots$. Then, there exists an equilibrium $(\mathbf{x}^*, \mathbf{y}^*)$ in $\Gamma_Q = \Gamma_{Q^1 Q^2}$ such that $(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ is a (completely mixed) equilibrium in Γ.*

Proof Take the sequences of strategies $\{\mathbf{x}_\ell^*\}_{\ell=1}^\infty \subset \mathbf{X}$ and $\{\mathbf{y}_\ell^*\}_{\ell=1}^\infty \subset \mathbf{Y}$ such that, for each $\ell = 1, 2, \ldots$, the strategy profile $(\mathbf{x}_\ell^*, \mathbf{y}_\ell^*)$ is the unique completely mixed equilibrium of Γ_{Q_ℓ}. Moreover, let $(\mathbf{x}^\dagger, \mathbf{y}^\dagger) \in \mathbf{E}^1 \times \mathbf{E}^2$ be the unique completely mixed equilibrium of Γ, which we know to be weakly completely mixed by Lemma 2. Then, applying Theorem 1, we have $\mathbf{x}_\ell^* Q_\ell^1 = \mathbf{x}^\dagger$ and $\mathbf{y}_\ell^* Q_\ell^2 = \mathbf{y}^\dagger$ for each $\ell = 1, 2, \ldots$.

Note that, because the strategy spaces $\mathbf{X}$ and $\mathbf{Y}$ are compact, there exists subsequences $\{\mathbf{x}_{\ell_s}^*\}_{s=1}^\infty$ and $\{\mathbf{y}_{\ell_t}^*\}_{t=1}^\infty$ that converge to some strategies $\mathbf{x}^* \in \mathbf{X}$ and $\mathbf{y}^* \in \mathbf{Y}$, respectively. Clearly,

$$\mathbf{x}^* Q^1 = \lim_{s \to \infty} \mathbf{x}_{\ell_s}^* Q_{\ell_s}^1 = \mathbf{x}^\dagger \quad \text{and} \quad \mathbf{y}^* Q^2 = \lim_{t \to \infty} \mathbf{y}_{\ell_t}^* Q_{\ell_t}^2 = \mathbf{y}^\dagger.$$

This shows that $(\mathbf{x}^* Q^1, \mathbf{y}^* Q^2)$ is a completely mixed equilibrium of Γ and, by Theorem 1(i), $(\mathbf{x}^*, \mathbf{y}^*)$ is an equilibrium of Γ_Q, as required. $\qquad\square$

2.4 Variational Properties

Now, we return to the dynamic setting where $\Gamma_Q(\lambda, \mu)$ denotes a family of incompetent games parameterised by a pair of learning trajectories. A central focus in the development of incompetence has been the variational properties of $\Gamma_Q(\lambda, \mu)$ when Γ is a matrix game or a bimatrix game. Here, we will summarise what is known about the behaviour of these incompetent games under variations in the players' learning parameters.

Beck *et* al. [9] study the dependence of equilibrium-induced expected rewards on the players' learning parameters. They present Theorem 3 and Theorem 4 showing that, under certain conditions on $Q^1(\lambda)$ and $Q^2(\mu)$, the expected rewards granted by a specific extreme equilibrium have useful representations.

Theorem 3 *[9] Assume that $Q^1(\lambda)$ and $Q^2(\mu)$ are linear, that is,*

$$Q^1(\lambda) = (1 - \lambda)Q^1(0) + \lambda Q^1(1) \quad and \quad Q^2(\mu) = (1 - \mu)Q^2(0) + \mu Q^2(1) \quad (14)$$

for all $\lambda, \mu \in [0, 1]$. *Fix* $\Lambda, M \subset [0, 1]$ *such that* $\Gamma_Q(\lambda, \mu)$ *share a* $h^1 \times h^2$ *Shapley-Snow kernel for all* $(\lambda, \mu) \in \Lambda \times M$. *Then, for some constants* $\alpha_{ij}^k, \beta_{ij}^k \in \mathbb{R}$, *the expected reward to Player* $k \in \{1, 2\}$ *achieved by the kernel's associated extreme equilibrium in* $\Gamma_Q(\lambda, \mu)$ *is*

$$\frac{\sum_{i=1}^{h^k+1} \sum_{j=1}^{h^k+1} \alpha_{ij}^k \lambda^{h^k-i+1} \mu^{h^k-j+1}}{\sum_{i=1}^{h^k} \sum_{j=1}^{h^k} \beta_{ij}^k \lambda^{h^k-i} \mu^{h^k-j}} \tag{15}$$

for each $(\lambda, \mu) \in \Lambda \times M$; *a ratio of bivariate polynomials in* λ *and* μ.

Theorem 4 *[9] Assume that* $Q^1(\lambda)$ *and* $Q^2(\mu)$ *are linear with initially uniform incompetence* $Q^1(0) = 1/m J_m$ *(or* $Q^2(0) = 1/n J_n$). *Then, the dependence of an extreme equilibrium's expected reward in (15) is (at most) linear in* λ *(or* μ).

Furthermore, in addition to proving specialisations of Theorem 3 and Theorem 4 for matrix games, Beck and Filar [10] establish several other properties regarding the game value of a parameterised incompetent matrix game $\Gamma_Q(\lambda, \mu)$. Specifically, they prove that the function $(\lambda, \mu) \mapsto \text{val}(\Gamma_Q(\lambda, \mu))$ is continuous and not-necessarily monotone in λ and μ. It is also shown that a player can never achieve a greater reward than under complete competence; that is,

$$\text{val}(\Gamma_{Q_1(\lambda), I_n}) \leq \text{val}(\Gamma_{Q_1(\lambda), Q_2(\mu)}) \leq \text{val}(\Gamma_{I_m, Q_2(\mu)}) \tag{16}$$

for all $\lambda, \mu \in [0, 1]$. Beck and Filar [10] also briefly address the plateauing game values of some parameterised incompetent matrix games (see, for example, Fig. 1) by noting that Corollary 1 might apply when a player approaches complete competence. The tools developed in Sect. 2.3 allow us to further explore this observation. Consider the set of learning parameters

$$\mathcal{C} := \{(\lambda, \mu) \in [0, 1] \times [0, 1] : \Gamma_Q(\lambda, \mu) \text{ is completely mixed}\} \tag{17}$$

on which $\Gamma_Q(\lambda, \mu)$ is completely mixed. Assume that the learning trajectories $Q^1(\lambda)$ and $Q^2(\mu)$ are continuous. Then, given that the set of reward matrices belonging to completely mixed matrix games is open (see Jansen [33]), the set $\mathcal{C}$ is also open. Theorem 2 shows that, for each $(\lambda, \mu) \in \overline{\mathcal{C}}$, the players are both able to execute a completely competent optimal strategy in $\Gamma_Q(\lambda, \mu)$. This means that, by an identical argument to Corollary 1, the function $(\lambda, \mu) \mapsto \text{val}(\Gamma_Q(\lambda, \mu))$ is constant on $\overline{\mathcal{C}}$. Hence, we expect a game value plateau to emerge whenever $\Gamma_Q(\lambda, \mu)$ becomes completely mixed.

2.5 Incremental Learning

Next, we will demonstrate a simple model of incremental learning in a parameterised family of incompetent matrix games $\Gamma_Q(\lambda, \mu)$. This *incremental learning game* Γ_{inc} is a stochastic game unfolding over an infinite time horizon $T = \{0, 1, 2, \ldots\}$ in which, between repeated plays of an incompetent game, the players may choose to increment their learning parameters through the ordered sets $\Lambda := \{\lambda_1, \lambda_2, \ldots, \lambda_M\}$ and $M := \{\mu_1, \mu_2, \ldots, \mu_N\}$. It is assumed that $\lambda_i < \lambda_{i+1}$ and $\mu_j < \mu_{j+1}$ for each $i = 1, 2, \ldots, M - 1$ and $j = 1, 2, \ldots, N - 1$. This means that a player's skill parameter can never be decreased or, informally, that a player can halt but never reverse the process of learning. Henceforth, we simplify notation by identifying i with λ_i and j with μ_j.
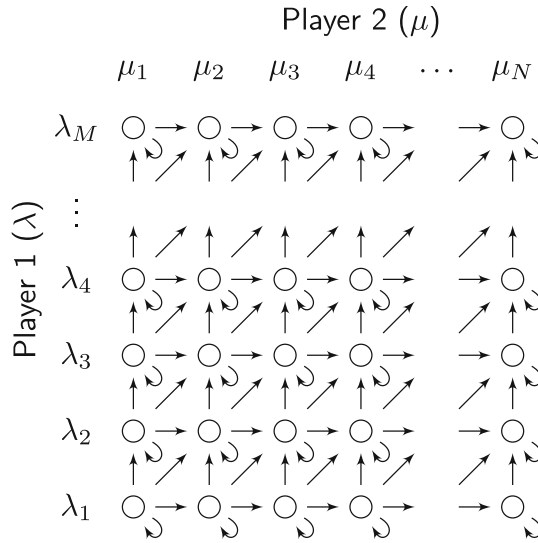
Fig. 3 The transition structure of a general incremental learning game, with nodes representing states and arcs representing potential transitions

Now, we give a precise description of Γ_{inc} using the language and notation associated with stochastic games in [18]. The state space

$$\mathcal{S} := \big\{(i, j) : i = 1, 2, \ldots, M \text{ and } j = 1, 2, \ldots, N \big\} \tag{18}$$

is chosen to index the learning parameters $\Lambda \times M$. Fix a stage $t \in T$ and a state $s = (i, j) \in \mathcal{S}$ such that λ_i and μ_j are the learning parameters belonging to Player 1 and Player 2 at stage t.

Player 1 and Player 2 (optimally) play the incompetent game $\Gamma_Q(i, j)$ and are given the option to advance their learning parameters to $i + 1$ and $j + 1$, respectively. The decision to increment a learning parameter might incur a state-dependent learning cost $c^k(i, j)$ to Player $k \in \{1, 2\}$. Formally, we say that the actions belonging to Player 1 and Player 2 at state s are

$$\mathcal{A}(s) := \begin{cases} \{0, 1\}, & i \neq M, \\ \{0\}, & i = M, \end{cases} \quad \text{and} \quad \mathcal{B}(s) := \begin{cases} \{0, 1\}, & j \neq N, \\ \{0\}, & j = N, \end{cases} \tag{19}$$

where "0" means "Don't Learn" and "1" means "Learn". If Player 1 selects $a \in \mathcal{A}(s)$ and Player 2 selects $b \in \mathcal{B}(s)$, then they receive the stage-t immediate rewards

$$r^k(s, a, b) := \begin{cases} \text{val}\big(\Gamma_Q(i, j)\big) - a c^1(i, j), & k = 1, \\ -\text{val}\big(\Gamma_Q(i, j)\big) - b c^2(i, j), & k = 2, \end{cases} \tag{20}$$

where the $\text{val}(\Gamma_Q(i, j))$ term is the reward received after optimally playing $\Gamma_Q(i, j)$. Moreover, before the subsequent $(t + 1)^{\text{th}}$ stage, the game transition to the state $(i + a, j + b)$ with (degenerate) transition probabilities given by

$$p(s'|s, a, b) := \begin{cases} 1, & s' = s + (a, b), \\ 0, & s' \neq s + (a, b), \end{cases} \tag{21}$$

for every $s' \in \mathcal{S}$. The general transition structure of this game is shown in Fig. 3.

Here, we will focus on *stationary strategies*, which are represented as block row vectors $\mathbf{f} = (\mathbf{f}(s))_{s \in \mathcal{S}}$ for Player 1 and $\mathbf{g} = (\mathbf{g}(s))_{s \in \mathcal{S}}$ for Player 2. The block $\mathbf{f}(s) = (f(s,a))_{a \in \mathcal{A}(s)}$ stores the probability $f(s,a)$ of choosing action $a \in \mathcal{A}(s)$ and the block $\mathbf{g}(s) = (g(s,b))_{b \in \mathcal{B}(s)}$ stores the probability of choosing action $b \in \mathcal{B}(s)$. The sets of stationary strategies belonging to Player 1 and Player 2 are denoted by $\mathbf{F}$ and $\mathbf{G}$, respectively. The immediate rewards in (20) and the transition probabilities in (21) are extended to $\mathbf{F} \times \mathbf{G}$ by defining

$$r^k(s, \mathbf{f}, \mathbf{g}) := \sum_{a \in \mathcal{A}(s)} \sum_{b \in \mathcal{B}(s)} f(s,a) r^k(s,a,b) g(s,b), \tag{22}$$

and

$$p(s' | s, \mathbf{f}, \mathbf{g}) := \sum_{a \, \mathcal{A}(s)} \sum_{b \in \mathcal{B}(s)} f(s,a) p(s' | s, a, b) g(s,b), \tag{23}$$

for each $(\mathbf{f}, \mathbf{g}) \in \mathbf{F} \times \mathbf{G}$. If the stochastic process $\{S_t\}_{t=0}^\infty$ stores the state at each stage $t \in T$, then it becomes a Markov chain under the dynamics induced by a strategy profile $(\mathbf{f}, \mathbf{g}) \in \mathbf{F} \times \mathbf{G}$. We use $\mathbb{P}_{s\mathbf{fg}}$ and $\mathbb{E}_{s\mathbf{fg}}$ to denote probabilities and expectations under these dynamics with the initial state $S_0 = s \in \mathcal{S}$. The β-discounted value ($\beta \in [0, 1)$) of $(\mathbf{f}, \mathbf{g}) \in \mathbf{F} \times \mathbf{G}$ to Player $k \in \{1, 2\}$ with the initial state $s \in \mathcal{S}$ is

$$v^k(s, \mathbf{f}, \mathbf{g}) := \sum_{t=0}^\infty \beta^t \mathbb{E}_{s\mathbf{fg}} \big[r^k(S_t, \mathbf{f}, \mathbf{g}) \big]. \tag{24}$$

Then, $(\mathbf{f}^*, \mathbf{g}^*) \in \mathbf{F} \times \mathbf{G}$ is a (Nash) equilibrium of the incremental learning game Γ_{inc} whenever

$$v^1(s, \mathbf{f}, \mathbf{g}^*) \leq v^1(s, \mathbf{f}^*, \mathbf{g}^*) \quad \text{and} \quad v^2(s, \mathbf{f}^*, \mathbf{g}) \leq v^2(s, \mathbf{f}^*, \mathbf{g}^*) \tag{25}$$

for all $s \in \mathcal{S}$, $\mathbf{f} \in \mathbf{F}$, and $\mathbf{g} \in \mathbf{G}$.

Although Γ_{inc} unfolds over an infinite time horizon, its transition structure admits a specialised backward induction algorithm for computing equilibria. We construct a suitable notion of "past" and "future" states by finding a sequence $s_1, s_2, \ldots, s_L$ (where $L := MN$) such that $\ell' < \ell$ implies $p(s_{\ell'} | \ell, a, b) = 0$ for all distinct $\ell, \ell' = 1, 2, \ldots, L$ and $(a, b) \in \mathcal{A}(s_\ell) \times \mathcal{B}(s_\ell)$.

It is straightforward to verify that a suitable ordering exists—for example, the lexicographical ordering. So, we shall assume that an ordering has been fixed and write ℓ instead of s_ℓ. Lemma 3 shows that the discounted value of a strategy profile at a specific state does not depend on the "past" states. This allows us to restrict the stochastic game Γ_{inc} to the limited state space $\{\ell, \ell + 1, \ldots, L\}$ while still being able to assess the value of strategies.

Lemma 3 *Fix* $(\mathbf{f}, \mathbf{g}) \in \mathbf{F} \times \mathbf{G}$. *Then, for any* $\ell \in \{1, 2, \ldots, L\}$ *and* $k \in \{1, 2\}$, *we have*

$$v^k(\ell, \mathbf{f}, \mathbf{g}) = \frac{r^k(\ell, \mathbf{f}, \mathbf{g}) + \beta \sum_{\ell'=\ell+1}^L v^k(\ell', \mathbf{f}, \mathbf{g}) p(\ell' | \ell, \mathbf{f}, \mathbf{g})}{1 - \beta p(\ell | \ell, \mathbf{f}, \mathbf{g})}. \tag{26}$$

Proof Observe that, by conditioning on the state S_1 after the first transition, the discounted value of $(\mathbf{f}, \mathbf{g})$ is

$$
v^k(\ell, \mathbf{f}, \mathbf{g}) = \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{\ell \mathbf{fg}}\big[r^k(S_t, \mathbf{f}, \mathbf{g})\big] = \mathbb{E}_{\ell \mathbf{fg}}\big[r^k(S_0, \mathbf{f}, \mathbf{g})\big]
$$

$$
+ \sum_{t=1}^{\infty} \sum_{\ell'=1}^{L} \beta^t \mathbb{E}_{\ell \mathbf{fg}}\big[r^k(S_t, \mathbf{f}, \mathbf{g})\big|S_1 = \ell'\big] \mathbb{P}_{\ell \mathbf{fg}}(S_1 = \ell')
$$

$$
\overset{*}{=} r^k(\ell, \mathbf{f}, \mathbf{g}) + \beta \sum_{\ell'=1}^{L} p(\ell'|\ell, \mathbf{f}, \mathbf{g}) \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{\ell' \mathbf{fg}}\big[r^k(S_t, \mathbf{f}, \mathbf{g})\big]
$$

$$
\overset{**}{=} r^k(\ell, \mathbf{f}, \mathbf{g}) + \beta \sum_{\ell'=\ell}^{L} p(\ell'|\ell, \mathbf{f}, \mathbf{g}) v^k(\ell', \mathbf{f}, \mathbf{g}).
$$

Note that the above equality $\overset{*}{=}$ can be verified by applying the definition of $r^k(\ell, \mathbf{f}, \mathbf{g})$ and appealing to the fact that $\{r^k(S_t, \mathbf{f}, \mathbf{g})\}_{t=0}^{\infty}$ is a Markov chain. Similarly, the equality $\overset{**}{=}$ holds by applying the definition of $v^k(\ell', \mathbf{f}, \mathbf{g})$. We now easily obtain (26) by rearranging to isolate the $v^k(\ell, \mathbf{f}, \mathbf{g})$ term on the left-hand side. □

Next, we will show that a backward induction algorithm can solve this incremental learning game by working backward through the states $1, 2, \ldots, L$. Fix a state $\ell \in \{1, 2, \ldots, L - 1\}$ and define $\mathbf{F}_{\ell+1} := \{(\mathbf{f}(\ell'))_{\ell'=\ell+1}^{L}\}$ and $\mathbf{G}_{\ell+1} := \{(\mathbf{g}(\ell'))_{\ell'=\ell+1}^{L}\}$ to be the sets of stationary strategies over the "future" states $\ell + 1, \ldots, L$. Assume that we have already found $\mathbf{f}_{\ell+1}^* \in \mathbf{F}_{\ell+1}$ and $\mathbf{g}_{\ell+1}^* \in \mathbf{G}_{\ell+1}$ solving (25) at each state $\ell' = \ell + 1, \ldots, L$. Can we extend this equilibrium to include the current state? Theorem 5 shows that it is sufficient to consider a simplified version of the equilibrium inequalities that only account for unilateral deviations at state ℓ. The sets $\mathbf{F}_{\ell}(\mathbf{f}_{\ell+1}^*) := \{(\mathbf{f}(\ell), \mathbf{f}^*(\ell + 1), \ldots, \mathbf{f}^*(L))\}$ and $\mathbf{G}_{\ell}(\mathbf{f}_{\ell+1}^*) := \{(\mathbf{g}(\ell), \mathbf{g}^*(\ell + 1), \ldots, \mathbf{g}^*(L))\}$ denote the spaces of stationary strategies that extend $\mathbf{f}_{\ell+1}^*$ and $\mathbf{g}_{\ell+1}^*$ to state ℓ.

Theorem 5 *The strategy profile* $(\mathbf{f}_{\ell}^*, \mathbf{g}_{\ell}^*) \in \mathbf{F}_{\ell}(\mathbf{f}_{\ell+1}^*) \times \mathbf{G}_{\ell}(\mathbf{g}_{\ell+1}^*)$ *is an equilibrium of* Γ_{inc} *restricted to* $\{\ell, \ldots, L\}$ *whenever*

$$
v^1(\ell, \mathbf{f}_{\ell}', \mathbf{g}_{\ell}^*) \le v^1(\ell, \mathbf{f}_{\ell}^*, \mathbf{g}_{\ell}^*) \quad \textit{and} \quad v^2(\ell, \mathbf{f}_{\ell}^*, \mathbf{g}_{\ell}') \le v^2(\ell, \mathbf{f}_{\ell}^*, \mathbf{g}_{\ell}^*) \tag{27}
$$

for all $\mathbf{f}_{\ell}' \in \mathbf{F}_{\ell}(\mathbf{f}_{\ell+1}^*)$ *and* $\mathbf{g}_{\ell}' \in \mathbf{G}_{\ell}(\mathbf{g}_{\ell+1}^*)$.

Proof We need to show that $(\mathbf{f}_{\ell}^*, \mathbf{g}_{\ell}^*)$ satisfying (27) also satisfies (25) at state ℓ. Take a pair of strategy profiles $(\mathbf{f}_{\ell}, \mathbf{g}_{\ell}) \in \mathbf{F}_{\ell} \times \mathbf{G}_{\ell}$ and $(\mathbf{f}_{\ell}', \mathbf{g}_{\ell}') \in \mathbf{F}_{\ell}(\mathbf{f}_{\ell+1}^*) \times \mathbf{G}_{\ell}(\mathbf{g}_{\ell+1}^*)$ such that $\mathbf{f}(\ell) = \mathbf{f}'(\ell)$ and $\mathbf{g}(\ell) = \mathbf{g}'(\ell)$. This means that $\mathbf{f}_{\ell}'$ (or $\mathbf{g}_{\ell}'$) is a combination of $\mathbf{f}_{\ell}$ (or $\mathbf{g}_{\ell}$) at ℓ and $\mathbf{f}_{\ell}^*$ (or $\mathbf{g}_{\ell}^*$) at $\ell + 1, \ldots, L$. So, $(\mathbf{f}_{\ell}', \mathbf{g}_{\ell})$ and $(\mathbf{f}_{\ell}', \mathbf{g}_{\ell}^*)$ satisfy the equilibrium inequalities at the states $\ell + 1, \ldots, L$; that is, we have

$$
v^1(\ell', \mathbf{f}_{\ell}, \mathbf{g}_{\ell}') \le v^1(\ell', \mathbf{f}_{\ell}', \mathbf{g}_{\ell}') \quad \text{and} \quad v^1(\ell', \mathbf{f}_{\ell}, \mathbf{g}_{\ell}^*) \le v^1(\ell', \mathbf{f}_{\ell}', \mathbf{g}_{\ell}^*)
$$

for any $\ell' \in \{\ell + 1, \dots, L\}$. Alongside the discounted value representation from Lemma 3, this gives

$$
\begin{aligned}
v^1(\ell, \mathbf{f}_\ell, \mathbf{g}_\ell^*) &= \frac{r^1(\ell, \mathbf{f}_\ell, \mathbf{g}_\ell^*) + \beta \sum_{\ell'=\ell+1}^{L} v^1(\ell', \mathbf{f}_\ell, \mathbf{g}_\ell^*) p(\ell'|\ell, \mathbf{f}_\ell, \mathbf{g}_\ell^*)}{1 - \beta p(\ell|\ell, \mathbf{f}_\ell, \mathbf{g}_\ell^*)} \\
&\leq \frac{r^1(\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*) + \beta \sum_{\ell'=\ell+1}^{L} v^1(\ell', \mathbf{f}_\ell', \mathbf{g}_\ell^*) p(\ell'|\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*)}{1 - \beta p(\ell|\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*)} = v^1(\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*)
\end{aligned}
$$

where $r^1(\ell, \mathbf{f}, \mathbf{g}_\ell^*) = r^1(\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*)$ and $p(\cdot|\ell, \mathbf{f}_\ell, \mathbf{g}_\ell^*) = p(\cdot|\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*)$ because $\mathbf{f}(\ell) = \mathbf{f}'(\ell)$. Finally, by (27), we obtain

$$
v^1(\ell, \mathbf{f}_\ell, \mathbf{g}_\ell^*) \leq v^1(\ell, \mathbf{f}_\ell', \mathbf{g}_\ell^*) \leq v^1(\ell, \mathbf{f}_\ell^*, \mathbf{g}_\ell^*).
$$

After repeating an analogous argument for Player 2, we see that the conditions in (27) are sufficient to ensure that that $(\mathbf{f}_\ell^*, \mathbf{g}_\ell^*)$ is an equilibrium of Γ_{inc} restricted to $\{\ell, \dots, L\}$. □

The useful consequence of Theorem 5 is that, by solving a "local" problem at "previous" state ℓ, we can extend the equilibrium $(\mathbf{f}_{\ell+1}^*, \mathbf{g}_{\ell+1}^*)$ to create $(\mathbf{f}_\ell^*, \mathbf{g}_\ell^*)$. This local problem resembles a repeated game with absorbing states. Namely, if the players both choose to forego learning, then the game remains at state ℓ. Otherwise, if either of the players choose to learn, then the game transitions into a new state where the expected future rewards are fixed by $(\mathbf{f}_{\ell+1}^*, \mathbf{g}_{\ell+1}^*)$. The rewards given to Player $k \in \{1, 2\}$ in this repeated game with absorbing states are

$$
V_{ab}^k = \begin{cases} r^k(\ell, a, b) + \beta v^k(s_\ell + (a, b), \mathbf{f}_{\ell+1}^*, \mathbf{g}_{\ell+1}^*), & (a, b) \neq (0, 0), \\ r^k(\ell, a, b), & (a, b) = (0, 0), \end{cases} \tag{28}
$$

for each $(a, b) \in \mathcal{A}(\ell) \times \mathcal{B}(\ell)$. An immediate consequence of Lemma 3 is that, for each $k \in \{1, 2\}$ and $(\mathbf{f}_\ell, \mathbf{g}_\ell) \in \mathbf{F}_\ell(\mathbf{f}_{\ell+1}^*) \times \mathbf{G}_\ell(\mathbf{g}_{\ell+1}^*)$, we have

$$
v^k(\ell, \mathbf{f}_\ell, \mathbf{g}_\ell) = \frac{1}{1 - \beta p_0 q_0} \sum_{a \in \mathcal{A}(\ell)} \sum_{b \in \mathcal{B}(\ell)} p_a V_{ab}^k q_b, \tag{29}
$$

where $p_0 = 1 - p_1 = f(\ell, 0)$ and $q_0 = 1 - q_1 = g(\ell, 0)$. Hence, to ensure that $(\mathbf{f}_\ell^*, \mathbf{g}_\ell^*) \in \mathbf{F}_\ell(\mathbf{f}_{\ell+1}^*) \times \mathbf{G}_\ell(\mathbf{g}_{\ell+1}^*)$ satisfies the inequalities in (27), we need to solve the coupled pair of maximisation problems

$$
\begin{cases} p_0^* = \underset{p_0 \in [0, 1]}{\arg\max} \dfrac{1}{1 - \beta p_0 q_0^*} \displaystyle\sum_{a \in \mathcal{A}(\ell)} \sum_{b \in \mathcal{B}(\ell)} p_a V_{ab}^1 q_b^*, \\[2ex] q_0^* = \underset{q_0 \in [0, 1]}{\arg\max} \dfrac{1}{1 - \beta p_0^* q_0} \displaystyle\sum_{a \in \mathcal{A}(\ell)} \sum_{b \in \mathcal{B}(\ell)} p_a^* V_{ab}^2 q_b, \end{cases} \tag{30}
$$

where $p_0^* = 1 - p_0^* = f^*(\ell, 0)$ and $q_0^* = 1 - q_1^* = g^*(\ell, 0)$. Under the additional assumption that this repeated game with absorbing states is non-degenerate, the solutions are either both pure strategies ($p_0^*, q_0^* \in \{0, 1\}$) or both completely mixed strategies ($p_0^*, q_0^* \in (0, 1)$). The pure strategy solutions can be found by imposing restriction $p_0, p_0^*, q_0, q_0^* \in \{0, 1\}$ in (30); that is, by comparing the payoffs of every possible pure strategy profile. Moreover, by setting the appropriate partial derivatives (with respect to p_0 and q_0, respectively)
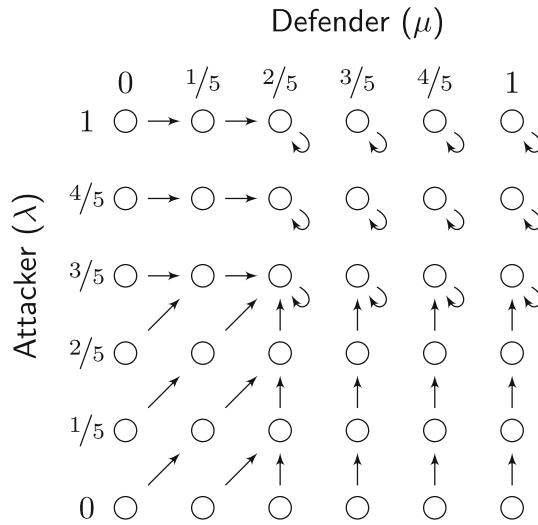
Fig. 4 The (unique) equilibrium learning strategies for the attacking and defending pilots with arrows showing the learning direction

of the functions being maximised in (30) equal to zero, we obtain

$$
\begin{cases}
\displaystyle\sum_{a\in\mathcal{A}(\ell)} p_a^* V_{a0}^2 - p_a^*(1-\beta p_0^*)V_{a1}^2 & = 0, \\
\displaystyle\sum_{b\in\mathcal{B}(\ell)} q_b^* V_{0b}^1 - q_b^*(1-\beta q_0^*)V_{1b}^1 & = 0.
\end{cases}
\tag{31}
$$

The solutions to (31) with $p_0^*, q_0^* \in (0,1)$ give the completely mixed strategy solutions to Γ_{inc}. This shows that we are always able to extend $(\mathbf{f}_{\ell+1}^*, \mathbf{g}_{\ell+1}^*)$ to an equilibrium $(\mathbf{f}_\ell^*, \mathbf{g}_\ell^*)$ of Γ_{inc} restricted to $\ell, \ldots, L$. Hence, since $(\mathbf{f}_L^*, \mathbf{g}_L^*)$ where $f^*(L,0) = g^*(L,0) = 1$ is the only strategy profile available at state L, we can work backwards through the states $L-1, L-2, \ldots, 2, 1$ and repeatedly extend it until obtaining an equilibrium $(\mathbf{f}^*, \mathbf{g}^*)$ of Γ_{inc}.

Example (attack-defence game with incremental learning) Lastly, recalling the attack-defence game $\Gamma_Q(\lambda, \mu)$ previously introduced in Sect. 2.2, suppose that the attacking and defending pilots have the option to undergo navigation training between engagements. We might model this as an incremental learning game Γ_{inc} in which training allows the pilots to advance their skill parameters through

$$
\Lambda = \{0, 1/5, 2/5, 3/5, 4/5, 1\} \quad \text{and} \quad M = \{0, 1/5, 2/5, 3/5, 4/5, 1\}
$$

after paying learning costs of $c^1(i,j) = c^2(i,j) = 1/10$ at state $s = (i,j) \in \mathcal{S}$. Moreover, assume that the pilots have far-sighted discounted strategy valuations with a discount factor of $\beta = 99/100$. What are the best strategies to reduce incompetence throughout this game?

The aforementioned backward induction algorithm produces a unique equilibrium of Γ_{inc} shown graphically in Fig. 4. A node indicates a pair of learning parameters and an arc indicates a transition realised by the equilibrium. So, a vertical arrow means that only Player 1 learns, a horizontal arrow means that only Player 2 learns, a diagonal arrow means that both players learn, and a loop means that neither player learns.

Note that, under the equilibrium shown in Fig. 4, the attacker learns until their skill reaches the interval $[^{11}/_{47}, 1]$ and the defender learns until they reach the interval $[^{26}/_{47}, 1]$. We know that the underlying parametrised incompetent game $\Gamma_Q(\lambda, \mu)$ is completely mixed on $(^{11}/_{47}, 1] \times (^{26}/_{47}, 1]$. So, by the observations in Sect. 2.4, both players are able to execute completely competent optimal strategies when $(\lambda, \mu) \in [^{11}/_{47}, 1] \times [^{26}/_{47}, 1]$. This means that, once the players have achieved learning parameters within these intervals, the game value plateaus (see Fig. .1) and there is no incentive to learn further. Therefore, it is not always necessary to achieve complete competence so long as the players are able to "mimic" competence by executing an optimal strategy from the completely competent game Γ.

3 Incompetence in Biological Populations

Game theory as a mathematical paradigm found applications not only in economics and behavioural studies, but also in biology. Its first application to biology was driven by the puzzling fact that animal contests rarely result in fights or serious injuries, even though contestants are sufficiently equipped to engage in an open fight [68]. It was suggested that instead of considering individuals as players who may not be rational, the selection itself could be considered as a rational force of evolution, and survival of the entire population is more important than benefits to individual members. Since then, evolutionary game theory emerged as a branch of game theory and ecological sciences studying evolution under selection pressure [28,50,56].

Recently, the effects of environmental changes on the evolution of biological populations became one of the main foci of the field [3,26,75,81]. Since all organisms on this planet live in a dynamic environment that undergoes changes, the ability to adapt becomes key to survival. Adaptation is a process that improves survival skills and reproductive functions of species, and usually includes two components: genetic adaptation and learning. As a specific example, when a population migrates or their environmental conditions change, their responses to new environmental stimuli may differ, introducing behavioural mistakes in individuals' interactions. The concept of incompetence was proposed in [37] to address the learning aspect of the evolution of social behaviour. Under the assumption of incompetence of individuals, behaviours that were likely to be observed in the old environment, might not have the same frequency in the new environment, and as organisms adapt, they might re-learn their previous behaviours.

3.1 Evolutionary Games

Naturally, game assumptions in biological settings differ from the classic games since rationality of each individual behaviour might not always be natural to assume. Consider a population of species consisting of N individual organisms. At every time step, individuals interact in a pair-wise manner, where they have to choose one action out of n distinct available actions. Outcomes of these interactions determine fitness of individuals based on the fitness matrix $R \in \mathbb{R}^{n \times n}$. In the evolutionary settings, all individuals in the population obtain the same fitness matrix R, however, during the interaction, Player 1's fitness is determined by R, while Player 2's fitness is determined by R^T. Furthermore, the sets of selectable and executable actions coincide for all players. Let $\mathbf{x} = (x_1, \ldots, x_n)$, where x_i denotes the frequency of the (pure) strategy i. We assume that in a given population, all individuals have

the same set of selectable actions $\mathcal{A}$, fitness matrix R, and the mixed strategy of the entire population $\mathbf{x}$.

The main focus of evolutionary games is to predict the strategy $\mathbf{x}$ that will be adopted by the population. Since we assume that n actions are available to each individual, the resulting mixed strategies lie in the simplex Δ_n defined by

$$\Delta_n = \left\{ \mathbf{x} = (x_1, \ldots, x_n) \middle| \sum_{i=1}^{n} x_i = 1, x_i \geq 0, \ \forall i = 1, \ldots, n \right\},$$

where $x_i = \frac{N_i}{N}$ with N_i being a number of individuals adopting strategy i and N being a total number of individuals in the population. Then, an evolutionary game Γ^e can be denoted by

$$\Gamma^e = \left\{ R, \mathcal{A}, \mathbf{x} \in \Delta_n \right\}. \tag{32}$$

We say that the population adopts a pure i^{th} strategy if all individuals are behaving as the ith type and, hence, their behavioural frequency vector is the unit basis vector $\mathbf{e}_i$. However, this may not always be the case. If not, we are in the case of mixed strategies $\mathbf{x}$, and hence we are interested in finding a mixture $\mathbf{x}^*$ which is a stable outcome of the evolution.

It was shown, that the concept of Nash equilibria is not sufficient when taking into account the evolution of populations [67]. As a result, a new equilibrium concept was proposed. The *evolutionary stable strategy (ESS)* ensures that population's strategy is resistant against random mutations and is defined, more precisely, below.

Definition 1 A mixed strategy $\mathbf{x}^*$ is called an evolutionary stable strategy if one of the following conditions hold:

(i) $\mathbf{x}^* R(\mathbf{x}^*)^T > \mathbf{y} R(\mathbf{x}^*)^T, \ \forall \mathbf{y} \in \Delta_n$;
(ii) if $\mathbf{x}^* R(\mathbf{x}^*)^T = \mathbf{y} R(\mathbf{x}^*)^T$, then $\mathbf{x}^* R \mathbf{y}^T > \mathbf{y} R \mathbf{y}^T, \ \forall \mathbf{y} \in \Delta_n$.

Here, $\mathbf{x}^* R(\mathbf{x}^*)^T$ measures the frequency-dependent fitness of the entire population, given that everyone adopts strategy $\mathbf{x}$, whereas $\mathbf{y} R(\mathbf{x}^*)^T$ measure fitness of a population adopting strategy $\mathbf{y}$ in a population of individuals using strategy $\mathbf{x}$. In the long run, an ESS guarantees that selection prefers $\mathbf{x}^*$ to any other arising strategy. Note that the ESS is a special case of a Nash equilibrium [56].

However, besides equilibria, we are usually interested in how these equilibria can be reached, bringing us to the concept of evolutionary dynamics. Given that biological populations not only interact, but also reproduce, there is a need to take into account the reproduction process. The first classic evolutionary dynamics model was proposed by Taylor and Jonker in [74], and is called *replicator dynamics*. These dynamics assume well-mixed infinitely large populations which is, of course, a simplification. Subsequently, many new concepts of dynamics were suggested in order to capture mutations [12,57,70,72], finite size of populations and stochasticity [31,53,73,76–78], adaptation [13,14,17,25,58], and a population structure [5,49,59,60]. However, to date, the concept of incompetence was only considered in a classic setting of replicator dynamics. In Conclusions section, we discuss possible extensions to other forms of dynamics for incompetent games.

Replicator dynamics captures a frequency-dependent selection, where the evolution of population's strategy depends on the current frequencies of all strategies in the population. That is, the fitness of a particular strategy is compared to the mean fitness of the entire population and is determined by the adopted strategies. With respect to a mixed strategy

$\mathbf{x} \in \Delta_n$, the expected fitness of a (pure) strategy i is defined by

$$f_i = \sum_{j=1}^{n} x_j r_{ij} = \mathbf{e}_i R \mathbf{x}^T = (R\mathbf{x})_i. \tag{33}$$

The mean fitness payoff of the population is then defined by the scalar

$$\phi = \sum_{i=1}^{n} x_i f_i = \mathbf{x} R \mathbf{x}^T. \tag{34}$$

Then, the dynamics of strategy i's frequency in the population is defined by

$$\dot{x}_i = x_i (f_i - \phi), \ i = 1, \dots, n,$$

or in a matrix form,

$$\dot{x}_i = x_i \left(\left(R\mathbf{x}^T \right)_i - \mathbf{x} R \mathbf{x}^T \right), \ i = 1, \dots, n. \tag{35}$$

In the folk theorem of evolutionary game theory, it was shown that any equilibrium of the replicator dynamics is a Nash equilibrium of the game Γ^e and that a strict Nash equilibrium is asymptotically stable [28]. Moreover, any ESS is an asymptotically stable equilibrium of the replicator dynamics. Hence, when considering evolutionary games, it is frequently sufficient to find equilibria of a static game Γ^e. This simplification is useful when trying to predict how the behaviour of the game changes under the assumption that interacting individuals are incompetent. We shall next consider how incompetence changes the game setup.

3.2 Evolutionary Games under Incompetence

When introducing an assumption that individuals are prone to making behavioural mistakes in an evolutionary game, one can interpret such mistakes as a form of behavioural plasticity. In some ways, this can be seen as phenotypic plasticity (for instance, in microbes). However, in application to more sophisticated organisms, behavioural plasticity need not relate to genetic background of the organism. These behavioural mistakes can be driven by migration to a new environment or any other form of environmental change and are reflected in the incompetence matrix Q analogous to that introduced in Sect. 2.2.

Since we assume that the entire population obtains only one fitness matrix, we also assume that the incompetence matrix is given for the entire population. Then, a new incompetent fitness matrix is determined in a similar manner to (7) as

$$R_Q = QRQ^T. \tag{36}$$

In line with previous sections, we assume that players' ability for improving their strategy execution is determined by some parameter. Since here we consider one population of players all of whom obtain the same measure of incompetence, we only need one incompetence parameter $\lambda \in [0, 1]$. Then, the incompetent fitness matrix is defined as

$$R(\lambda) := R_{Q(\lambda)} = Q(\lambda) R Q(\lambda)^T. \tag{37}$$

Throughout this section, we make a specific assumption on the functional form of learning. We assume that $Q(\lambda)$ is linear and defined as

$$Q(\lambda) = (1 - \lambda)S + \lambda I, \tag{38}$$

where S is the staring level of incompetence and I is the identity matrix. When $\lambda = 1$, the population does not make any execution errors and has a perfect strategy execution. Now we can define the evolutionary incompetent game as

$$\Gamma_Q^e = \left\{ R, \mathcal{A}, \mathbf{x} \in \Delta_n, Q(\lambda) : \lambda \in [0, 1] \right\}. \tag{39}$$

We can further simplify the analysis by utilising the property of replicator dynamics that it is invariant under a linear positive transformation [27]. This allows us to reduce the fitness matrix by subtracting diagonal elements of R from the corresponding columns. Mathematically speaking, such transformation can be defined as

$$\tilde{R} := R - \mathbf{d}_R \mathbf{1}_n^T, \tag{40}$$

where $\mathbf{d}_R$ is a vector consisting of the diagonal elements of R and $\mathbf{1}_n$ is a vector consisting of ones. Throughout the manuscript, we shall denote any matrix $\tilde{R}$ as a canonical form of matrix R as in (40). Then, according to (33)-(35), for a new game under incompetence Γ_Q^e, we re-write the expected fitness for strategy i as

$$f_i(\lambda) = \sum_{j=1}^n \tilde{r}_{ij}(\lambda) x_j = \mathbf{e}_i \tilde{R}(\lambda) \mathbf{x}^T, \tag{41}$$

and for the mean fitness payoff of the population,

$$\phi(\lambda) = \sum_{i=1}^n x_i f_i(\lambda) = \mathbf{x} \tilde{R}(\lambda) \mathbf{x}^T. \tag{42}$$

Hence, the incompetent replicator dynamics can be written as

$$\dot{x}_i = x_i(f_i(\lambda) - \phi(\lambda)), \ i = 1, \dots, n. \tag{43}$$

In a strict sense, the new system given by (43) is a perturbed evolutionary game, and perturbations depend on the parameter λ. As λ tends to 1 for all i, the game under incompetence approaches the original game given by R. In the following section, we summarise the main results obtained for incompetent evolutionary games.

3.3 Equilibria Transitions

Here, we are mostly interested in behaviours that dynamics exhibit under changes in parameter values λ given the starting level of incompetence S and the fitness matrix R. These behaviours may arise for different values of λ and the dynamics change their behaviour at critical levels λ^c of λ, referred to as bifurcation points, where equilibria emerge, disappear or change their stability properties.

Definition 2 [37] *A critical value λ^c of the incompetence parameter* is the bifurcation point of the replicator dynamics.

Under incompetence, behaviour of the game dynamics may exhibit several bifurcations [37]. Since by design the incompetence parameter approaches 1 when incompetence decreases, the incompetent fitness matrix $R(\lambda)$ is approaching the original fitness matrix R. As a result, in the limit of perfect competence, behaviour of the incompetent game approaches the behaviour of the original game. That is, there exists a maximal critical value of λ, that preserves robust properties of the game. We recall this result in the following theorem.

Theorem 6 *[37] If the game $\tilde{R}$ possesses an ESS, $\mathbf{x}^*$, and $||Q(\lambda) - I|| \leq \delta(\lambda^u)$, where $\lambda^u = \max \lambda^c$ is the maximal critical value of the incompetence parameter for a fixed point $\mathbf{x}^*$, then the incompetent game $\tilde{R}(\lambda)$, when $\lambda \in (\lambda^u, 1]$, possesses an ESS, $\mathbf{x}^*(\lambda)$, and*

$$\lim_{\lambda \to 1^-} \mathbf{x}^*(\lambda) = \mathbf{x}^*. \tag{44}$$

A natural question arises of how these bifurcation values of the incompetence parameter can be determined and the behaviour of dynamics. The larger the game (the more available strategies it has), the harder it becomes to define all possible bifurcations. However, even for an arbitrary number of strategies, we can find bifurcations of special equilibria, such as, interior equilibria or pure-strategies equilibria using analysis presented in [11]. Let us first focus on the bifurcations of the interior equilibria.

Definition 3 *[37]* Let $\mathbf{x}^*$ be a fixed point and λ^c be a bifurcation point that is also a zero of the mean fitness, namely, $\phi(\mathbf{x}^*, \lambda^c) = 0$. Then, λ^c is *a balanced bifurcation parameter value.*

Then, the point of bifurcation for an interior equilibrium can be found by considering a determinant of the incompetent fitness matrix.

Lemma 4 *[37] If $\mathbf{x}^*$ is an interior fixed point, that is, $x_i^* > 0, \forall i$. Then every balanced bifurcation parameter value, λ^c, is also a singular point of $\tilde{R}(\lambda)$ in the sense that $\det(\tilde{R}(\lambda)) = 0$.*

Next, we recall the special canonical form of the matrix $\tilde{R}(\lambda)$ that is defined through a rank-one transformation of an incompetent fitness matrix $R(\lambda)$. By [24], its determinant can be written as

$$\det(\tilde{R}(\lambda)) = \det(R(\lambda) - \mathbf{d}_{R(\lambda)} \mathbf{1}_n^T) = (1 - \mathbf{1}_n^T R(\lambda)^{-1} \mathbf{d}_{R(\lambda)}) \det(R)[\det(Q(\lambda))]^2. \tag{45}$$

Hence, critical values of the incompetence parameter can be found by finding zeroes of either $\det(Q(\lambda))$, or $[1 - \mathbf{1}_n^T R(\lambda)^{-1} \mathbf{d}_{R(\lambda)}]$.

In a special case of a rock-paper-scissors game [40], stability of the interior equilibrium is determined by the sign of the determinant of the fitness matrix [80], which gives rise to three cases: (a) if $\det(R) < 0$, then an unstable interior equilibrium exists resulting in a heteroclinic cycle; (b) if $\det(R) > 0$, then such an equilibrium is a stable mixed equilibrium; (c) if $\det(R) = 0$, then there exists a centre and periodic orbits around it.

However, games $\tilde{R}(\lambda)$ and $R(\lambda)$ exhibit the same behaviour [27]. Since the determinant of the fitness matrix $R(\lambda)$ always preserves the same sign as $\det(R)$, then $\det(\tilde{R}(\lambda))$ also cannot change its sign while the interior equilibrium exists.

Deriving a general form of equilibria depending on λ is complex and depends on the form of matrices R and S. For a special case of uniform incompetence, which implies that everybody makes mistakes with the same probability $1/n$, we can sometimes find a closed-form expression for the interior equilibrium. A uniform incompetence can be interpreted as a form of plasticity in biological populations. For instance, phenotypic plasticity, when different types might have slight variations in the exact degree of each gene expression. We provide this result in the following theorem.

Theorem 7 *[40] Let $\mathbf{x}^*$ be an interior ESS for R. Then, for λ sufficiently close to 0, if the starting level of incompetence, S, is a uniform matrix, that is, $s_{ij} = 1/n, \forall i, j = 1, \ldots, n,$ then*

$$\mathbf{x}^*(\lambda) = \frac{1}{\lambda} \left(\mathbf{x}^* - \frac{1 - \lambda}{n} \mathbf{1}_n \right) \tag{46}$$

is an interior ESS for the game $\tilde{R}(\lambda)$.

In [11], it was shown that pure-strategy equilibrium's stability properties can be determined from the sign of the j-th column of matrix $\tilde{R}(\lambda)$. Hence, given the maximal level of incompetence, we can determine which of the vertices will be stable and when this stability will change.

Theorem 8 *[40] If*

$$(\mathbf{s}_l - \mathbf{s}_j)^T R \mathbf{s}_j < 0, \ \forall \, l \neq j$$

then vertex j is a stable point of the replicator dynamics with execution errors for $\lambda \in [0, \lambda^c)$, where λ^c is the smallest critical value of the incompetence parameter where $\tilde{r}_{lj}(\lambda_{lj}^c)$ changes its sign for some $l \neq j$.

This result can be generalised for any level of incompetence, where we will have to consider

$$(\mathbf{q}_l - \mathbf{q}_j)^T R \mathbf{q}_j < 0, \ \forall \, l \neq j,$$

for all levels of λ, where $\mathbf{q}_i$ is the ith row of $Q(\lambda)$. Generally speaking, this condition implies that for a pure strategy to be stable under incompetence, it is necessary for it to be the best response to itself given all other pure strategies.

As in [38], in Fig. 5, we provide some cases illustrating how equilibria stability can change as competence level of the population changes, for an example of an unstable rock-paper-scissors game. In every panel, the colour-coded bar at the top indicates which of the three vertices is stable, while in the main plot we depict the interior equilibrium components as functions of λ. Even in this simple game, the behaviour exhibited by the replicator dynamics in response to changing λ can be very rich. As shown in these three examples, the interior equilibrium may or may not exist for different values of the incompetence parameter. Similarly, one, two, three or none of the vertices may be stable at the same time.

As in the case of classical games, in evolutionary settings it is natural to consider decreasing levels of incompetence, a process we called *learning*. Note that increasing λ corresponds to greater skill level and decreasing incompetence.

In the evolutionary games setup, dynamic incompetence was interpreted from two different perspectives: as an environmental shift that requires adaptation from organisms before the stable equilibrium is reached and as a learning process designed to maximise the fitness of the population after the population stabilised at some equilibrium.

Behaviour exhibited by the population dynamics when the process of learning is treated as a function of time, $\lambda(t)$, was considered in [36,38]. There are many options possible when choosing the functional form. So far, two functional forms of $\lambda(t)$ were analysed: a sigmoid and a periodic function. The sigmoid form of learning implies that organisms are capable of learning faster in the beginning of the process and slower when they reach sufficient competence. An assumption of a slowing rate for high enough levels of λ is motivated by the absence of necessity to learn fast since the evolutionary stable outcome can already be reached (see Theorem 6).

Analysing parameters of $\lambda(t)$ one can determine how long will it take for species to be fully recovered in behavioural sense and act as in the environment they are familiar with. However, while the functional form of the adaptation trajectory captures the pace and steepness of the learning process, the starting level of incompetence can be seen as a measure of the magnitude of the changes in the environment. That is, the further the new habitat is from the previous one, the longer it may take for organisms to fully recover.

Since natural habitats are prone to some form of regular stochasticity, in [36], it was also considered how periodic environmental fluctuations due to the seasonal or daily changes affect
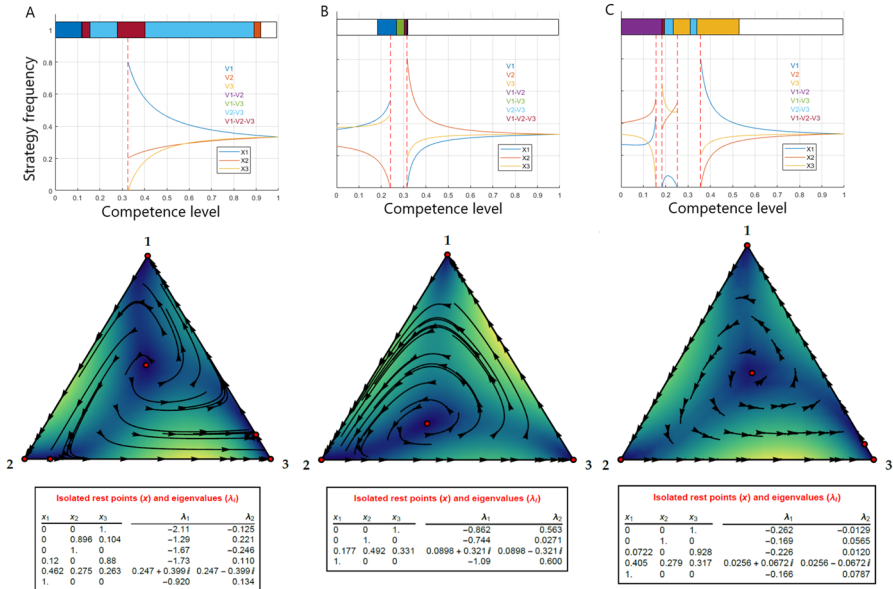
Fig. 5 Examples of equilibria transitions for different cases of an unstable rock-paper-scissors game. Here, we consider three examples with different matrices S_i and R_i in each panel A, B and C. First, we plot the interior equilibria transitions for $\lambda \in [0, 1]$ and report intervals of λ where one, two or all three vertices become stable in a colour-coded bar. Then, for $\lambda = 1/2$ we plot a phase diagram in Δ_3 of the corresponding replicator dynamics [32]. In the table for each example we report eigenvalues of the Jacobian of the replicator dynamics for each of the equilibria of the dynamics for a given $\lambda = 1/2$. Interested reader is referred to [38] for the code used to produce plots in A, B and C.

the evolutionary dynamics. It appeared that periodicity of environmental changes leads to periodic behaviour in the evolutionary dynamics as well. Specifically, if the original game possesses a stable equilibrium, then the solution of the incompetent game with periodic form of incompetence will converge to a stable periodic orbit around this stable equilibrium.

Let us now demonstrate how the concept of incompetence can be applied to a more specific biological setting. In the next section, we formulate a game of two foraging strategies of marine bacteria and try to analyse it from the perspective of incompetence.

3.4 Bacterial Motility Game under Incompetence

Evolutionary game theory has been widely applied to studying the evolution of microbes. Despite their primitivism and small sizes, marine bacteria are among the most ubiquitous forms of marine organisms, playing a central role in governing health of marine ecosystems and regulating global biosphere [52]. Understanding how cells make decisions and interact has implications for both biology of bacterial communities and our exploitation of these communities [1,20,21,44,46,48]. Fundamental to nutrient competition among bacteria is the choice of motility (chemotactic) strategies. Chemotaxis—the ability to sense environmental signals, and react to the stimuli accordingly—has been studied since the late 1800s [16,61].

However, a deterministic game theoretic approach misses an essential feature of the bacterial population dynamics: these populations and their interactions are highly stochastic. For instance, stochastic environmental fluctuations often affect ecological systems [21]. In order

to at least partially allow for this, the concept of incompetence was applied to study foraging strategies of bacteria in [36] by incorporating behavioural stochasticity in a matrix game that captures interactions between different strategic types of microbes. The aim is to identify the most efficient strategy for given environmental conditions. We consider two possible strategies: nonmotile or chemotactic. Nonmotile bacteria cannot induce active swimming and only drift with the water flow, whereas chemotaxis allows for active choice of direction. The fitness matrix can be constructed as

$$
\begin{array}{cc}
 & \text{Nonmotile} \quad \text{Chemotactic} \\
\begin{array}{c} Nonmotile \\ \text{Chemotactic} \end{array} & \begin{pmatrix} 1 & 0 \\ 1 - c + m & \frac{1}{2} - c \end{pmatrix},
\end{array}
$$

where c is the cost of swimming and m is the reward for being able to efficiently determine the direction of swimming and both parameters are normalised so that $c, m \in [0, 1]$. Depending on the exact values of the parameters, the game might exhibit four different behaviours, as proposed in [82] in relation to the signs of matrix elements in a canonical form from (40):

1. *Nonmotile strategy dominates:* for $c > \frac{1}{2}$ and $m < c$;
2. *Chemotactic strategy dominates:* for $c < \frac{1}{2}$ and $m > c$;
3. *A stable mixed equilibrium exists:* for $c > \frac{1}{2}$ and $m > c$;
4. *An unstable mixed equilibrium exists:* for $c < \frac{1}{2}$ and $m < c$.

We shall focus on two cases: when chemotactic strategy dominates nonmotile strategy and when a stable mixed equilibrium exists. The mixed equilibrium is given by

$$
x_N = \frac{2(m - c)}{2m - 1} \quad \text{and} \quad x_C = \frac{2c - 1}{2m - 1}. \tag{47}
$$

When introducing incompetence in a model, one has to take into account biological limitations of the strategies. For instance, there exist no conditions under which a nonmotile bacterium can exhibit chemotaxis because it lacks receptors and flagella required for such a strategy. However, a chemotactic bacterium can be both nonmotile and chemotactic. Hence, the starting incompetence matrix S for this example, may have the following form

$$
S = \begin{array}{c} \text{Nonmotile} \\ \text{Chemotactic} \end{array} \begin{array}{cc} \text{Nonmotile} & \text{Chemotactic} \\ \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \end{array}
$$

Then, the resulting induced incompetent fitness matrix $\tilde{R}(\lambda)$ is given by

$$
\tilde{R}(\lambda) = \begin{array}{c} \text{Nonmotile} \\ \text{Chemotactic} \end{array} \begin{array}{cc} \text{Nonmotile} & \text{Chemotactic} \\ \begin{pmatrix} 0 & \frac{1}{4}\left(4c + 2m - 1 + \lambda(2m - 1)\right) \\ m - c & 0 \end{pmatrix}. \end{array}
$$

Note that the relative fitness of chemotactic strategy is not affected by incompetence. However, the advantage of nonmotile strategy depends on the level of incompetence induced by chemotactic bacteria. By using Lemma 4, we can determine the critical value of the
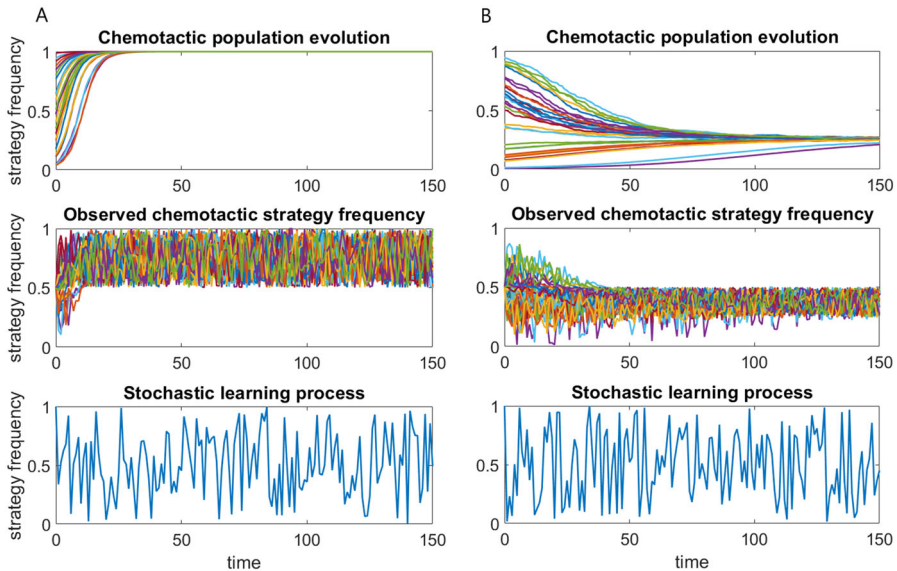
Fig. 6 Stochastic incompetence in the bacterial motility game: **a** for $m = 0.75$ and $c = 0.25$, **b** for $m = 0.75$ and $c = 0.7$. In the first row we plot 100 simulated solutions of the replicator dynamics of the chemotactic strategy, $x_C(\lambda, t)$, for a given stochastic process of $\lambda(t)$. In the second row we depict the executed or observed frequency of chemotactic strategy as a function of t and λ. The corresponding stochastic learning process $\lambda(t)$ is depicted in the last row of Fig.

incompetence parameter, which is determined as the solution of $\det(\tilde{R}(\lambda)) = 0$ or $\tilde{r}_{21}(\lambda^c) = 0$, and given by

$$\lambda^c = \frac{1 - 4c + 2m}{2m - 1}.$$

Depending on the stability properties of the dynamics in the original game, the behaviour of equilibria under incompetence will differ. For instance, if chemotactic strategy was dominating in a fully competent game, then for $\lambda < \lambda^c$ both strategies will stably co-exist. If a stable mixed equilibrium existed, then for $\lambda < \lambda^c$ chemotactic strategy will dominate nonmotile strategy.

Additionally, turbulence affects life of marine bacteria [71]. Mathematically, this can be modelled via a stochastic adaptation process. Construct a stochastic learning process where each point $\lambda(t)$ is a random variable with distribution that is determined by the species' migration process. This assumption provides a more realistic interpretation of the species' behaviour when we take into account migration and environmental stochasticity. However, games with an ESS are well-known to be robust [11].

We shall compare population dynamics for two types of the learning processes: deterministic sigmoid learning and a stochastic migration process. Any learning process, either deterministic or stochastic, will lead to the ESS in terms of a species' choice (Fig. 6). Observations of the real behaviour depend on the incompetence matrix. Population's behavioural observations for different learning processes will be different depending on the learning dynamics.

The stochastic learning brings us to a situation where the majority of bacteria in a population are able to perform chemotaxis if chemotactic strategy was dominating (Fig. 6, left), as

in the deterministic case. However, if there exists a stable mixed equilibrium, then dynamics under stochastic incompetence converges to the stable frequency of chemotactic bacteria as well (Fig. 6, right). This is yet different from the equilibrium of the game without incompetence.

Due to incompetence, extinct strategies may still reappear in the behaviour of individuals as a manifestation of mistakes that cause a revival of the extinct types. This randomisation may become beneficial as a changing environment may require flexibility from individuals in their adaptation.

Even if the adaptive peak has been reached (i.e. $\lambda = 1$), behavioural randomisation may become essential in preparedness to unforeseen changes. This is supported by the existing research in stochastic phenotype switching, when bacteria perform behavioural stochasticity even in stable environments [30].

When considering incompetent games, the main focus of the analysis is on where the dynamics will stabilise and whether a stable equilibrium will be reached. However, what if the change in the environment happened after the stable equilibrium was already reached? Is there an optimal way to re-learn effective strategies that is least costly in terms of fitness losses? We discuss results answering this question in the next section.

3.5 Prioritised Learning

When allowing for learning after the stable equilibrium is reached, the focus of the analysis is the population's need to re-learn its effective strategies in an optimal manner. Hence, in [39] the learning under incompetence was considered with respect to maximising the fitness over the learning path.

When addressing learning, one needs to distinguish whether the entire population is learning with the rate λ or whether each strategy has its own learning rate $\lambda_i \in [0, 1]$. This decision depends on the specific situation under consideration. For this section, let us assume a more general case with $\lambda = (\lambda_1, \ldots, \lambda_n)$ to define an evolutionary game under incompetence. Then, a performance measure of the learning path over fitness can be thought of as

$$\Phi_{C^*}(\lambda) = \max_C \int \phi_C(\lambda) d\lambda,$$

where C is a learning path that can be taken and $\phi_C(\lambda)$ is the mean-fitness of the population. Note that here it is explicitly assumed that every strategy i has its own incompetence parameter λ_i, which implies that $\lambda = (\lambda_1, \ldots, \lambda_n)$. Since the complexity of the problem grows with the number of strategies, this model was considered in its simplest possible setup: when only two strategies compete. That is, the fitness matrix R has the canonical form

$$
\begin{array}{cc}
 & \begin{array}{cc} \text{Strategy 1} & \text{Strategy 2} \end{array} \\
\begin{array}{c} \text{Strategy 1} \\ \text{Strategy 2} \end{array} & \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}
\end{array}
$$

and the starting level of incompetence S can be denoted as

$$S = \begin{pmatrix} \eta & 1-\eta \\ 1-\gamma & \gamma \end{pmatrix}.$$

If the initial game has one stable pure equilibrium, then the optimal learning will simply imply reduction of frequency of execution of the unwanted strategy. However, if the game

Table 1 Definitions of advantages of Strategy 1 over Strategy 2. For the definition of advantages of Strategy 2 over Strategy 1, the inequality signs in parameters comparison should be reversed

Advantage	Parameters comparison	Discussion
Fitness advantage	$a > b$	Fitness advantage implies that Strategy 1 has higher fitness and, hence, is more abundant.
Learning advantage	$\eta < \gamma$	Learning advantage implies that Strategy 1 is more flexible in its execution of Strategies 1 and 2.
Strategic advantage	$\delta > 0$ or $\tilde{a} > \tilde{b}$	Strategic advantage combines both fitness and learning advantages implying that if Strategy 1 is disadvantaged in fitness (or learning), then advantage in learning (or fitness) can compensate.

possesses a mixed stable equilibrium, it is no longer obvious what the learning path should look like. Note that an interior equilibrium in a 2-strategy game has the following form $\hat{\mathbf{p}} = (\hat{a}, \hat{b})$, where

$$\hat{a} := \frac{a}{a+b} \quad \text{and} \quad \hat{b} := \frac{b}{a+b}. \tag{48}$$

In order to maximise the fitness of the population, it is sufficient to consider the mean fitness function [74], which has the following form

$$\phi = \hat{\mathbf{p}} R \hat{\mathbf{p}}^T = \frac{ab}{a+b},$$

which under incompetence is reduced to the analysis of two parameters

$$\tilde{a} := \frac{\eta - \hat{a}}{1 - \eta}, \quad \tilde{b} := \frac{\gamma - \hat{b}}{1 - \gamma}. \tag{49}$$

An important aspect stemming from this model is the understanding of fitness and learning advantages. Given that relative fitness of each strategy is positive, that is, $a, b > 0$, we say that the strategy with higher relative fitness obtains a *fitness advantage*.

In addition, we say that strategies may obtain a *learning advantage*. This concept is induced by incompetence and implies that the strategy with more variability in the behaviour, that is, with a higher probability of mistakes, has a higher potential fitness advantage which might be achieved by reducing incompetence. Hence, the lower η or γ, the greater the learning advantage. A new parameter $\delta := \tilde{a} - \tilde{b}$ was defined to measure the *relative strategic advantage* of one strategy over another. We summarise and compare these concepts in Table 1.

Hence, since we allow for one strategy to have a relative strategic advantage over another one, the optimal learning path depends on which strategy is advantageous. This phenomenon was called *prioritised learning*.

Definition 4 We say that there exists *prioritised learning* for $\Phi_C(\lambda)$ among stepwise learning paths, if there exists C^* such that one of the directions is preferable over the other. That is, $\Phi_{C_1}(\lambda) \neq \Phi_{C_2}(\lambda)$, where $\Phi(\lambda)$ is the fitness-over-learning depending on the direction of learning i and C_1, C_2 are the learning paths in directions 1 and 2, respectively.

Interestingly, the sign of δ fully determines which strategy has to be learnt first. That is, it cannot be determined separately based on either fitness or learning advantages of strategies. The naive suggestion would be that the most advantageous skill in terms of fitness has to be learnt first. However, the strategy with lower relative strategic advantage is learnt first in the optimal learning path.

Theorem 9 *[39] The direction of the optimal learning path is determined by the sign of δ: for $\delta > 0$ the direction of Strategy 2 is optimal and for $\delta < 0$ the direction of Strategy 1 is optimal. If $\delta = 0$, then there is no difference in the direction of optimal learning, that is, $\Phi_{C_1}(\lambda) = \Phi_{C_2}(\lambda)$.*

We suggest that natural selection tries to compensate the most disrupted strategy first even if its fitness is not the highest. Nonetheless, if the fitness difference is high enough to overcome the effect of incompetence, then the optimal learning will demand that the better strategy is learned first. Another possible interpretation would be to consider the mixed equilibrium as mixed strategies used by players. Then, by learning the less-advantageous strategy first, individuals are reaching the nearest optimal mixed strategy.

In the next section we shall demonstrate results from three previous sections on a reduced 2-strategy game based on the foraging strategies of marine bacteria as presented in [36].

3.6 Bacterial Motility Game and Prioritised Learning

Let us now assume that the population has stabilised at the mixed equilibrium defined in (47). Assume that the environmental conditions have changed leading to deviations in strategy executions for both bacterial strategies. For this, assume that the new starting incompetence matrix is defined as

$$
\hat{S} = \begin{array}{c} \\ \text{Nonmotile} \\ \\ \text{Chemotactic} \end{array} \begin{array}{cc} \text{Nonmotile} & \text{Chemotactic} \\ \left(\begin{array}{cc} 1 - \epsilon_1 & \epsilon_1 \\ \\ \epsilon_2 & 1 - \epsilon_2 \end{array} \right). \end{array}
$$

Since nonmotile bacteria can exhibit chemotactic behaviour only as a random noise, it is natural to assume that $\epsilon_1 \leq \epsilon_2$. Furthermore, let us allow for each strategy to be learnt at a different pace as in Sect. 3.5. In order to determine the optimal path that maximises the fitness over learning, we first calculate advantages of nonmotile and chemotactic strategies from (49) as

$$
\tilde{a} = \frac{1 - \epsilon_1}{\epsilon_1} + \frac{2m - 2c}{\epsilon_2(2m - 1)} \quad \text{and} \quad \tilde{b} = \frac{1 - \epsilon_2}{\epsilon_2} + \frac{2c - 1}{\epsilon_2(2m - 1)}.
$$

Then, the strategic advantage of nonmotile strategy over chemotactic strategy equals to

$$
\delta = \left(\frac{1 - \epsilon_1}{\epsilon_1} - \frac{1 - \epsilon_2}{\epsilon_2} \right) + \left(\frac{2m - 2c}{\epsilon_2(2m - 1)} - \frac{2c - 1}{\epsilon_2(2m - 1)} \right).
$$

Note that if $\epsilon_1 = \epsilon_2 = \epsilon$, then $\delta = 1/\epsilon > 0$ and chemotactic strategy has to be learnt first in one step (see Fig. 7 (left)). Generally, chemotactic strategy has to be learnt first whenever $\delta > 0$ or equivalently
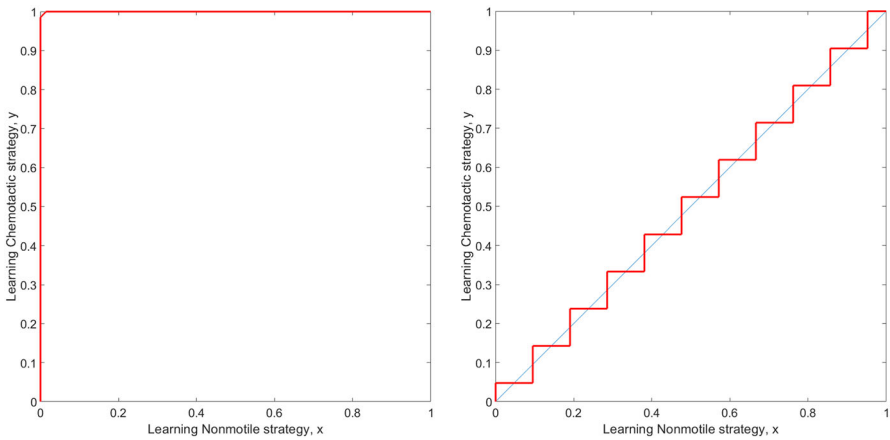
$$
\frac{\epsilon_1}{\epsilon_2} > \frac{2m - 2c}{2c - 1},
$$

Fig. 7 Prioritised learning paths for the bacterial motility game for $m = 0.9$, $c = 0.75$ and: (left) $\epsilon_1 = \epsilon_2 = 0.1$, (right) $\epsilon_1 = 1/10$ and $\epsilon_2 = 1/6$

which together with the condition $\epsilon_1 < \epsilon_2$ requires that $m < 2c - \frac{1}{2}$. We plot a special case when $\delta = 0$ in Fig. 7 (right).

4 Conclusions and Future Extensions

This paper is predicated on the belief that competitions/games with incompetent agents/players are ubiquitous in nature. Hence, formalising the notion of incompetence and modelling the impact of the resulting "mistakes" on the outcomes of games is worthy of detailed analysis. However, we first must recognise that everyday use of the word "incompetence" carries a very wide range of possible interpretations and hence needs to be narrowed down in order to be rigorously analysed.

Hence, the line of research we surveyed is limited to situations where incompetence can be adequately modelled via probability distributions on specified sets of actions available to one or more players (assumption [A1]). This implies that incompetence induced mistakes manifest themselves as random outcomes, different from intended outcomes. The latter certainly captures some essential characteristics of incompetence.

However, in the case of classical incompetent games studied so far, assumption [A1] was augmented by a requirement that players know one another's propensity to make mistakes. This "mutually known" aspect concerning the probability distributions of mistaken executions is certainly restrictive. For instance, it is clear that while it may approximately apply to a match between two professional tennis players at Wimbledon, it would not hold for two children playing one another. We hope that future investigations will relax this restriction.

4.1 Extensions for Classical Games

Currently, in the setting of classical nooncooperative game theory, incompetence has been studied mainly in matrix and bimatrix games. However, there are clearly several other, more general, classes of games to which this approach could be extended. Below, we name just four, out of many possible, generalisations.

a) *Continuum of actions.* Although we have only dealt with players having finitely-many actions, the concept of incompetence could be extended to games with larger action spaces, for example, games with a continuum of actions. Given a game with a continuum of actions, mixed strategies are represented by cumulative distribution functions and expected utility is computed as a Riemann-Stieltjes integral. This means that, in this context, a general "incompetence-adjusted" utility function (as in (8)) would also need to be expressed in an integral form.

b) *Incompetence dependent action spaces.* While the original incompetence framework described by Beck *e*t al. [9] allows a player's selectable and executable actions to differ, the theoretical development to date addressed only the case where they coincide. Intuitively, it is clear that there are situations where a player's incompetence may contract or expand their set of selectable actions. However, this raises the conceptual challenge of dynamically capturing the changes to these sets, as a player reduces his or her incompetence via learning. This would need to be modelled in a sufficiently general and yet technically tractable way.

c) *Extensions to stochastic games.* In stochastic games evolving over discrete time horizon, at each stage players play one of a finite set non-cooperative games called "states". The consequence of a single play is an immediate payoff (to each player) and a probabilistic transition to a new state (e.g. see the seminal paper [65]). Clearly, it is possible to replace each state by an incompetent non-cooperative game, thereby inducing an incompetent stochastic game. Such a generalisation would be interesting and, likely, tractable.

d) *Extensions to incremental learning.* The incremental learning games formulated in Sect. 2.5 adopt several simplifying assumptions that could be relaxed to further extend the model. First, the assumption that a player's learning trajectory can be parameterised by a single learning parameter could be relaxed to allow for "multidirectional learning". Second, relaxing the assumption that a player's level of incompetence can never be decremented would allow the model to describe not only the process of learning, but also the process of forgetting what one has learnt.

4.2 Extensions for Evolutionary Games

It should be clear that the work done so far in studying incompetent evolutionary games constitutes merely a beginning. As above, in this section we briefly describe just three, out of many, possible continuations of this research.

a) *Generalisations of population dynamics.* First of all, choosing to work within replicator dynamics setting carries with it simplifying assumptions which open it to criticism for oversimplification of natural reproduction processes. While replicator dynamics is a classical approach to modelling the effect of natural selection, over decades of research, these assumptions were relaxed in new approaches to modelling population dynamics.

In particular, the effect of the finite population size and inherent stochasticity of the reproduction process were addressed in finite population dynamics like Moran birth-death process, ability of others to imitate successful behavioural aspects of neighbours were addressed in imitation dynamics and the effect of interaction with neighbours was addressed in many different dynamics on networks. Hence, as a natural extension, one should consider how relaxed assumptions on the population dynamics affect the dynamics of games with incompetence.

b) *Generalising prioritised learning by exploiting the power of simulations.* In recent years, evolutionary models adopted methods of computer simulations to facilitate exploration of more realistic complex models that would have been intractable using analytical methods.

Hence, one could extend our prioritised learning of Sect. 3.5 to allow more than two strategies to compete at the same time. Furthermore, it is tempting to allow every individual organism their own learning parameter so as to closer approximate natural scenarios.

While the complexity of such a complex model will render it intractable analytically, simulating specific setups may shed light on many puzzling biological problems. For instance, the problem of determining how niches emerge and are filled by organisms while interacting under many different environmental conditions with multiple organisms.

c) *Learning as a function of frequency of strategies.* One simplifying assumption we made so far in all models of incompetence applied in biology is that of the separation of the learning process from the reproduction process. However, evolution of learning or levels of incompetence might be frequency dependent, which could lead to intricate co-evolution. Setups that follow similar logic were considered in [81] and [75]. While results presented in [75] can be seen as more general, the form of the exact dynamics of the learning process in the settings of incompetence was not addressed in previous works. Hence, we believe it would be worthwhile to consider co-dependence of $\mathbf{x}(\lambda)$ and $\lambda(t, \mathbf{x})$.

Data Availability The data sets and simulations discussed in this article are available from the relevant co-author upon reasonable request.

References

1. Abrudan M, You L, Staňková K, Thuijsman F (2016) A game theoretical approach to microbial coexistence. In advances in dynamic and evolutionary games. Springer, Berlin, pp 267–282
2. Adami C, Hintze A (2018) Thermodynamics of evolutionary games. Phys Rev E 97(062136):1–8
3. Akçay E (2020) Deconstructing evolutionary game theory: coevolution of social behaviors with their evolutionary setting. Am Nat 195(2):315–330
4. Albrecht A, Avrachenkov K, Howlett P, Verma G (2020) Evolutionary dynamics in discrete time for the perturbed positive definite replicator equation. ANZIAM J 62:148–184
5. Allen B, Lippner G, Chen YT, Fotouhi B, Momeni N, Yau ST, Nowak MA (2017) Evolutionary dynamics on any population structure. Nature 544(7649):227–230
6. Avrachenkov K, Borkar VS (2019) Metastability in stochastic replicator dynamics. Dyn Games Appl 9(2):366–390
7. Beck J (2013) Incompetence, training and changing capabilities in game theory. Ph.D. thesis, University of South Australia, Australia
8. Beck JD (2011) Game theory implementation of capability investment problem. Mil Op Res 16(1):41–55
9. Beck JD, Ejov V, Filar JA (2012) Incompetence and impact of training in bimatrix games. Automatica 48(10):2400–2408
10. Beck JD, Filar JA (2007) Games, incompetence and training. Ann ISDG 8:93–110

11. Bomze I (1986) Non-cooperative two-person games in biology: a classification. Int J Game Theory 15:31–57
12. Bomze I, Burger R (1995) Stability by mutation in evolutionary games. Games Econ Behav 11(2):146–172
13. Dercole F, Rinaldi S (2008) Analysis of evolutionary processes: the adaptive dynamics approach and its applications. Princeton University Press, USA
14. Dieckmann U, Marrow P, Law R (1995) Evolutionary cycling in predator-prey interactions: population dynamics and the red queen. J Theor Biol 176(1):91–102
15. Dridi S (2019) Plasticity in evolutionary games. bioRxiv p. 509604
16. Engelmann TW (1883) Bacterium photometricum. Archiv für die gesamte Physiologie des Menschen und der Tiere 30(1):95–124
17. Eshel I (1983) Evolutionary and continuous stability. J Theor Biol 103(1):99–111
18. Filar JA, Vrieze K (1997) Competitive Markov Decision Processes. Springer, USA
19. Foster D, Young P (1990) Stochastic evolutionary game dynamics. Theor Popul Biol 38(2):219–232
20. Frey E (2010) Evolutionary game theory: theoretical concepts and applications to microbial communities. Physica A: Statist Mech Appl 389(20):4265–4298
21. Frey E, Reichenbach T (2011) Bacterial games. Principles of Evolution pp. 297–329
22. Fudenberg D, Harris C (1992) Evolutionary dynamics with aggregate shocks. J Econ Theor 57(2):420–441
23. Fudenberg D, Levine D (1999) The theory of learning in games. The MIT Press, USA
24. Harville D (1997) Matrix algebra from a statistician's perspective, vol 1. Springer, USA
25. Hilbe C, Abou Chakra M, Altrock PM, Traulsen A (2013) The evolution of strategic timing in collective-risk dilemmas. PloS One 8(6):e66490
26. Hilbe C, Schmid L, Tkadlec J, Chatterjee K, Nowak MA (2018) Indirect reciprocity with private, noisy and incomplete information. Proc Natl Acad Sci 115(48):12241–12246
27. Hofbauer J, Schuster P, Sigmund K, Wolff R (1980) Dynamical systems under constant organization ii: homogeneous growth functions of degree p=2. SIAM J Appl Math 38(2):282–304
28. Hofbauer J, Sigmund K (2003) Evolutionary game dynamics. Bullet Am Math Soc 40(3):479–519
29. Hopkins E (2002) Two competing models of how people learn in games. Econometrica 70(6):2141–2166
30. Hufton P, Lin Y, Galla T (2018) Phenotypic switching of populations of cells in a stochastic environment. J Statist Mech: Theor Exp 2018(2):023501
31. Imhof LA, Nowak MA (2006) Evolutionary game dynamics in a wright-fisher process. J Math Biol 52(5):667–681
32. Izquierdo LR, Izquierdo SS, Sandholm WH (2018) Evodyn-3s: a mathematica computable document to analyze evolutionary dynamics in 3-strategy games. SoftwareX 7:226–233
33. Jansen MJM (1981) Regularity and stability of equilibrium points of bimatrix games. Math Op Res 6(4):530–550
34. Jurg AP, Jansen MJM, Parthasarathy T, Tijs SH (1990) On weakly completely mixed bimatrix games. Linear Algebra Appl 141:61–74
35. Kaplansky I (1945) A contribution to von neumann's theory of games. Ann Math 46(3):474–479
36. Kleshnina M (2019) Evolutionary games under incompetence and foraging strategies of marine bacteria. Ph.D. thesis, The University of Queensland. PhD thesis
37. Kleshnina M, Filar JA, Ejov V, McKerral JC (2018) Evolutionary games under incompetence. J Math Biol 77(3):627–646
38. Kleshnina M, McKerral JC, Gonzalez-Tokman C, Filar JA, Mitchell JG (2020) Shifts in evolutionary balance of microbial phenotypes under environmental changes. bioRxiv
39. Kleshnina M, Streipert SS, Filar JA, Chatterjee K (2020) Prioritised learning in snowdrift-type games. Mathematics 8(11):1945
40. Kleshnina M, Streipert SS, Filar JA, Chatterjee K (2021) Mistakes can stabilise the dynamics of rock-paper-scissors games. PLOS Comput Biol 17(4):e1008523
41. Komarova N (2004) Replicator-mutator equation, universality property and population dynamics of learning. J Theor Biol 230:227–239
42. Komarova N, Niyogi P, Nowak M (2001) The evolutionary dynamics of grammar acquisition. J Theor Biol 209:43–59
43. Kuhn HW (1961) An algorithm for equilibrium points in bimatrix games. Proc Natl Acad Sci 47(10):1657–1662
44. Lambert G, Vyawahare S, Austin R (2014) Bacteria and game theory: the rise and fall of cooperation in spatially heterogeneous environments. Interface Focus 4(4):1–12
45. Larkey P, Kadane JB, Austin R, Zamir S (1997) Skill in games. Management Science **43**(5)
46. Lenski R, Velicer G (2000) Games microbes play. Selection 1(3):89–95
47. Levin S (2003) Complex adaptive systems: exploring the known, the unknown and the unknowable. Bullet Am Math Soc 40(1):3–19

48. Li XY, Pietschke C, Fraune S, Altrock P, Bosch T, Traulsen A (2015) Which games are growing bacterial populations playing? J Royal Soc Interface 12(108):1–10
49. Lieberman E, Hauert C, Nowak MA (2005) Evolutionary dynamics on graphs. Nature 433(7023):312–316
50. McKelvey R, Apaloo J (1995) The structure and evolution of competition-organized ecological communities. Rocky Mt J Math 25(1):417–436
51. McKelvey R, Palfrey T (1995) Quantal response equilibria for normal form games. Games Econ Behav 10:6–38
52. Mitchell J (1991) The influence of cell size on marine bacterial motility and energetics. Microb Ecol 22(1):227–238
53. Moran P, Alfred P et al. (1962) The statistical processes of evolutionary theory. The statistical processes of evolutionary theory
54. Nash J (1951) Non-cooperative games. Annals Math 54(1):286–295
55. von Neumann J (1928) Zur Theorie der Gesellschaftsspiele. Mathematische Annalen 100(1):295–320
56. Nowak M (2006) Evolutionary dynamics: exploring the equations of life. The Belknap press of Harvard University Press, UK
57. Nowak M, Komarova N, Niyogi P (2001) Evolution of universal grammar. Science 291(5501):114–118
58. Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological games. Science 303(5659):793–799
59. Nowak MA, Tarnita CE, Antal T (2010) Evolutionary dynamics in structured populations. Philos Trans Royal Soc B: Biol Sci 365(1537):19–30
60. Perc M, Gómez-Gardenes J, Szolnoki A, Floría LM, Moreno Y (2013) Evolutionary dynamics of group interactions on structured populations: a review. J Royal Soc Interface 10(80):20120997
61. Pfeffer W (1884) Locomotorische Richtungsbewegungen durch chemische Reize:(Aus den" Untersuchungen aus dem botanischen Institut zu Tübingen Bd. I. Heft 3 p. 363-482). W. Engelmann
62. Raghavan TES (1970) Completely mixed strategies in bimatrix games. J London Math Soc 2(4):709–712
63. Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. Int J Game Theory 4(1):25–55
64. Selten R (1991) Evolution, learning and economic behavior. Games Econ Behav 3:3–24
65. Shapley LS (1953) Stochastic games. Proc Natl Acad Sci 39:1095–1100
66. Shapley LS, Snow RN (1952) Basic solutions of discrete games. Princeton University Press, USA, pp 27–36
67. Smith J (1982) Evolution and the theory of games. Cambridge University Press, USA
68. Smith J, Price G (1973) The logic of animal conflict. Nature 246:15–18
69. Spencer H (1864) The principles of biology. 1: London. Edinburgh, Williams and Norgate, 474p
70. Stadler P, Schuster P (1992) Mutation in autocatalytic reaction networks. J Math Biol 30(6):597–631
71. Stocker R (2012) Marine microbes see a sea of gradients. Science 338(6107):628–633
72. Tarnita CE, Antal T, Nowak MA (2009) Mutation-selection equilibrium in games with mixed strategies. J Theor Biol 261(1):50–57
73. Taylor C, Fudenberg D, Sasaki A, Nowak MA (2004) Evolutionary game dynamics in finite populations. Bullet Math Biol 66(6):1621–1644
74. Taylor P, Jonker L (1978) Evolutionary stable strategies and game dynamics. Math Biosci 40:145–156
75. Tilman AR, Plotkin JB, Akçay E (2020) Evolutionary games with environmental feedbacks. Nat Commun 11(1):1–11
76. Traulsen A, Claussen JC, Hauert C (2005) Coevolutionary dynamics: from finite to infinite populations. Phys Rev Lett 95(23):238701
77. Traulsen A, Hauert C (2009) Stochastic evolutionary game dynamics. Rev Nonlin Dyn Complex 2:25–61
78. Traulsen A, Shoresh N, Nowak MA (2008) Analytical results for individual and group selection of any intensity. Bullet Math Biol 70(5):1410
79. Vorob'ev NN (1958) Equilibrium points in bimatrix games. Theor Probab Appl 3(3):297–309
80. Weissing F (1991) Evolutionary stability and dynamic stability in a class of evolutionary normal form games In Game Equilibrium Models I. Springer, Berlin, pp 29–97
81. Weitz JS, Eksin C, Paarporn K, Brown SP, Ratcliff WC (2016) An oscillating tragedy of the commons in replicator dynamics with game-environment feedback. Proc Natl Acad Sci 113(47):E7518–E7525
82. Zeeman E (1980) Population dynamics from game theory In global theory of dynamical systems. Springer, Berlin, pp 471–497