Large Deviations And Stochastic Stability In Population Games

Journal of Dynamics and Games
doi:10.3934/jdg.2021021

LARGE DEVIATIONS AND STOCHASTIC STABILITY IN POPULATION GAMES

Mathias Staudigl*

Maastricht University
Department of Data Science and Knowledge Engineering
Paul-Henri-Spaaklaan 1, 6229 EN Maastricht, The Netherlands

Srinivas Arigapudi

Technion-Israel Institute of Technology
Faculty of Industrial Engineering and Management
Haifa, 32000003, Israel

William H. Sandholm

University of Wisconsin
Department of Economics
1180 Observatory Drive, Madison, WI 53706, USA

(Communicated by Michel Benaïm)

Abstract. In this article we review a model of stochastic evolution under general noisy best-response protocols, allowing the probabilities of suboptimal choices to depend on their payoff consequences. We survey the methods developed by the authors which allow for a quantitative analysis of these stochastic evolutionary game dynamics. We start with a compact survey of techniques designed to study the long run behavior in the small noise double limit (SNDL). In this regime we let the noise level in agents' decision rules to approach zero, and then the population size is formally taken to infinity. This iterated limit strategy yields a family of deterministic optimal control problems which admit an explicit analysis in many instances. We then move in by describing the main steps to analyze stochastic evolutionary game dynamics in the large population double limit (LPDL). This regime refers to the iterated limit in which first the population size is taken to infinity and then the noise level in agents' decisions vanishes. The mathematical analysis of LPDL relies on a sample-path large deviations principle for a family of Markov chains on compact polyhedra. In this setting we formulate a set of conjectures and open problems which give a clear direction for future research activities.

1. **Introduction.**

Evolutionary game theory (EGT) is the study of strategic interactions involving a large number of agents under mean-field interactions. Opposite to traditional game theory, EGT emphasizes the importance of dynamic processes, which are meant as *macroscopic* descriptions of behavioral adjustment processes

2020 *Mathematics Subject Classification.* Primary: 91A22, 60F10, 60J10; Secondary: 35F21, 90B20.

Key words and phrases. Markov chains, large deviations, Hamilton-Jacobi equations, state-constrained optimal control.

M. Staudigl is supported by the COST Action CA-16228.Srinivas Arigapudi is supported in part at the Technion by a Fine Fellowship.

* Corresponding author: Mathias Staudigl.

over time. The key quantity of interest in EGT is the long-run evolution of the relative frequencies that some actions are visible in a large (technically a continuum) population. To set the stage explaining this, let us introduce the basic terminology prevalent in standard EGT models.

1.1. **Population games.** We start this overview article with a quick refresher on standard terminology on normal form games. There is a finite set of player roles $\mathcal{P} = \{1, 2, \dots, \bar{p}\}$ where $\bar{p} \geq 1$. Each player role is characterized by a finite set $S^p = \{1, 2, \dots, n_p\}$, which we shall call the set of pure actions. We shall assume that $n_p \geq 1$ for all p, with strict inequality for at least on p. Let $\Delta^p = \{\sigma^p \in \mathbb{R}^{n_p}_+ | \sum_{i\in S^p} \sigma^p(i) = 1\}$ the set of mixed strategies of player role p. A tuple $\sigma = (\sigma^p; \sigma^{-p})$ defines a point in product set $\Delta := \prod_{p\in\mathcal{P}} \Delta^p$, and is called a *mixed strategy profile*. The payoff function of player role p is a vector-valued map

$$g^p : \Delta \to \mathbb{R}^{n_p}, x \mapsto g^p(\sigma) = [g^p_i(\sigma)]_{i\in S^p}.$$

The real-valued function $g^p_i(x)$ is interpreted as the *expected* payoff of action $i \in S^p$, given the mixed strategy profile $\sigma \in \Delta$ is used by the players. Define

$$\bar{g}^p(\sigma) = \sum_{i\in S^p} \sigma^p_i g^p_i(\sigma),$$

the average payoff of player role p. The pure best response of player role p is defined as

$$b^p(\sigma) = \text{argmax}\{g^p_j(\sigma) | j \in S^p\},$$

and the mixed best response is defined as

$$B^p(\sigma) = \text{argmax}\{\langle \tau^p, g^p(\sigma)\rangle \,| \tau^p \in \Delta^p\}.$$

Definition 1.1. A mixed action profile σ is a *Nash equilibrium* if

$$\sigma^p \in B^p(\sigma) \qquad \forall p \in \mathcal{P}. \tag{1}$$

The set of Nash equilibria is denoted by $\Delta^{NE} \subset \Delta$.

The interpretation of Nash equilibrium is an old, partly philosophical, debate. A very compelling interpretation has been given by John Nash himself in what has been popularized as the *mass action interpretation* [36, 58]. Assume that there are N agents forming a large population. Before the game is played each agent is randomly assigned to a single player role $p \in \mathcal{P}$. This assignment is done once and for all. The overall mass of player role p is m^p, so that the $\bar{p}$ populations are of sizes $(Nm^1, \dots, Nm^{\bar{p}})$. Let $N^p = Nm^p$ the number of agents in population $p \in \mathcal{P}$. A population state x^p is an action distribution $x^p = (x^p_1, \dots, x^p_{n_p})$ contained in the set $\mathsf{X}^p = m^p\Delta^p$. A social state is a collection of strategy distributions $x = (x^1, \dots, x^{\bar{p}})$, living in the polyhedron $\mathsf{X} = \prod_{p\in\mathcal{P}} \mathsf{X}^p$. The population p payoff vector is defined as $F^p : \mathsf{X} \to \mathbb{R}^{n_p}$, where $F^p_i(x)$ is the payoff associated with the action $i \in S^p$. The collection of vector payoffs $F = (F^1, \dots, F^{\bar{p}})$ defines a *population game*. [43] is the authoritative reference on this model. For concreteness, we will henceforth focus on the simpler *single population case*, in which $\bar{p} = 1$. In this simplified setting, the population mass can be, without loss of generality, be normalized to 1. The population game will then be identified with the vector-valued mapping $F : \mathsf{X} \to \mathbb{R}^n$, where $\mathsf{X} = \Delta = \{x \in \mathbb{R}^n_+ | \sum_{i\in S} x_i = 1\}$. Concrete examples will be given in Section 2.

1.2. **Evolutionary game dynamics.** Traditionally, evolutionary game dynamics have been formulated using autonomous differential equations, reflecting different behavioral principles (including best-response, imitation and pairwise comparison). Differential equation models contain fundamental adjustment dynamics like the Replicator dynamics [33, 55, 26, 21], the Brown-von Neumann-Nash dynamics [23], and the best-response dynamics [22, 27]. This stream of literature has been fundamental to provide a dynamical systems foundation of Nash equilibrium. However, several celebrated impossibility results [20, 25] also showed that the dynamical aspects themselves reveal interesting properties of game dynamics, and not only the mono-centric focus on equilibrium. Soon it has been realized that stochastic perturbations of basic evolutionary game dynamics could lead to new interesting phenomena. The literature on *stochastic evolutionary game dynamics* (SEGT) has been pioneered by [17, 59], where the first fundamental definitions around *stochastic stability* have been presented. In this literature, stochastic stability was introduced as a new equilibrium notion. Inspired by the fundamental theorem of game theory, relating fixed points of deterministic flows to equilibria of the underlying game, a stochastically stable equilibrium was defined in terms of invariant measures of the stochastic evolutionary process. Key to this approach is the classical mathematical theory on perturbed Markov chains [18, 29, 28]. Peyton H. Young, in his pathbreaking article [59], brought the general apparatus of Markov chains with rare transitions to the attention of game theorists. Connections with the general theory of interacting particle systems have been first established in [8]. [57] applied the same tool in the analysis of stochastic games. Applications of stochastic evolutionary game dynamics to problems in engineering and economics can be found in [13, 12, 32, 50, 31, 49], to name just a few. Even more research activity on stochastic evolutionary game dynamics can be located within mathematical biology; See [56] for a survey.

Unfortunately, the analysis of learning dynamics in games building on the notion of stochastic stability remained at an infant stage for a long time. Beside potential games, and simple stochastic models, no general tool has been known which allowed researchers to compute stochastically stable equilibria. Additionally, it must be emphasized that stochastic stability is, by definition, a concept which aims to understand the asymptotic properties of a stochastic process. As such it deals with quantifying hitting times and the characterization of supports of invariant measures. The typical formulation in this field involves very large times scales, and the complexity of the process itself cannot be really understood with the methods developed in this field. The main reason for this is that stochastic stability estimates usually capture only asymptotics at an exponential scale and thus are not fine grained enough to provide accurate asymptotic expressions. This led many researchers applying stochastic evolutionary game dynamics to concrete problems to the conclusion that this equilibrium selection device is not practically relevant. We do not fully support this point of view. First, and foremost, stochastic evolutionary game dynamics do not only shed light in the way how a population eventually settles in equilibrium, but also gives quantitative insights how equilibrium breaks down (metastability). This is an additional layer, which is not visible in related deterministic approaches. Second, stochastic game dynamics are deeply related to other fundamental models in applied mathematics, including mean-field Markovian dynamics in statistical physics (e.g. Curie-Weiss models), evolutionary biology [56] and optimization and control [9, 38]. Second, in order to develop refined estimates about the long-run behavior of the stochastic learning dynamics, sample path large deviation tools are a very good guideline to get first rough insights into the process, and to help understanding what tools are needed to refine the analysis. This is the case when looking at mixing times (a topic not covered in this article), or general metastable phenomena [10].

Stochastic stability analysis in games is an application of large-deviations theory to a specific Markovian system. In particular, we are interested in quantifying large deviation properties of a sequence of Markov chains $\mathbf{X}^{N,\eta} = \{X_k^{N,\eta}\}_{k\in\mathbb{N}_0}$ characterized by two crucial parameters (N, η). N is the size of the system. In a statistical physics context, it would represent the number of particles. η is a scaling parameter of the random noise in the system. From a statistical physics point of view it corresponds to the temperature. Both parameters can be studied at their extreme values to obtain idealized limit predictions:

1. The *small noise limit* (SNL; N fixed and $\eta \to 0$): This is the classical stochastic stability scenario. One reconstructs the support of the limiting invariant measure with the help of the Markov chain tree theorem of [11, 18]. Key to this approach is the identification of the rate of decay of the transition probabilities characterizing the finite Markov chain. This led to the introduction of an *unlikelihood function*, a concept first systematically introduced in games in [42] and [53].
2. The *large population limit* (LPL; $N \to \infty$ and $\eta > 0$ fixed): This is the setting where deterministic differential equations provide a good description of the transient phase of the stochastic process. In this limit, the Markov chain can be seen as a particular stochastic approximation scheme of a system of first-order differential equations. For the analysis, one relies on the ODE approach of stochastic approximation [30, 4, 45].
3. *Small noise double limit* (SNDL; $\eta \to 0$ and then $N \to \infty$): This can be analyzed via optimal control techniques. This gives rough estimates on waiting times and stationary distribution weights (See [52] and [46]).
4. *Large population double limit* (LPDL; $N \to \infty$ and then $\eta \to 0$): This limit is largely unexplored. The only reported results in this limit case are in [42]. In this overview article we will try to spell out the main difficulties arising in this important limit case.

Remark 1. These sequential limits are, of course, just extreme ways how the limiting dynamics could be analyzed. It might be equally interesting to study limits where $(N, \eta_N) \to (\infty, 0)$ at some specified rate. We are not aware of any analysis in this regime.

The large population double limit is mathematically different from the conceptually simpler small noise double limit. The common characteristic of both limiting scenarios is that they can be described in terms of deterministic optimal control problems. However, these control problems share fundamentally different characteristics. While the small noise limit is in general characterized by solutions to differential inclusions, the large population double limit leads to a more standard description of a deterministic optimal control problem. Also the "zero-cost solutions" of these optimal control problems are different. While the zero cost paths in the small noise limit correspond to solutions of the *controllable best-response dynamics* [46], in the large population limit these are given by solutions of the mean-field dynamics of the stochastic process. In this article we summarize the key results established to obtain rough asymptotic estimates of key statistics of the population process. For the limits SNL, SNDL and LPL we give a concise summary in Section 4 and 5.2. LPDL is not understood yet, so we only make a list of conjectures illustrated with the help of a minimal example (Section 5.3) showing that in general this limit is complicated. Overall, we collect seven open problems, all of which would be important next milestones helping us to understand the long-run behavior of stochastic evolutionary game dynamics.

Notation. Let $\mathsf{X} \subset \mathbb{R}^n$ be a closed convex nonempty set. The tangent cone is defined as $\mathsf{T}_{\mathsf{X}}(x) = \text{cl}[\mathbb{R}_+(\mathsf{X} - x)]$. For $T \in (0, \infty)$, let $\mathbf{C}([0,T] : \mathsf{X})$ denote the set of continuous functions $\phi\colon [0,T] \to \mathsf{X}$ through X over time interval $[0,T]$, endowed with the supremum norm. Let $\mathbf{C}_x[0,T]$ denote the set of paths $\phi \in \mathbf{C}([0,T] : \mathsf{X})$ with initial condition $\phi_0 = x$, and let $\mathbf{AC}_x[0,T]$ be the set of absolutely continuous paths in $\mathbf{C}_x[0,T]$. For every $T > 0$ we measure distance between any two points $\phi, \psi \in \mathbf{C}([0,T] : \mathsf{X})$ by the supremum norm $\rho_T(\phi,\psi) = \max_{0 \leq t \leq T} \|\phi(t) - \psi(t)\|$, and $\|\cdot\|$ denotes the ℓ^1 distance on $\mathbb{R}^n$. For elements $\phi, \psi \in \mathbf{C}(\mathbb{R}_+ : \mathsf{X})$ we define the metric

$$\rho(\phi,\psi) = \int_0^\infty e^{-t}\left(\rho_t(\phi,\psi) \wedge 1\right) \mathrm{d}t,$$

which generates the topology of uniform convergence on compact sets on $\mathbf{C}(\mathbb{R}_+ : \mathsf{X})$.

We let, $\mathbb{R}^n_0 = \{x \in \mathbb{R}^n | x_1 + x_2 + \cdots + x_n = 0\}$. $\Delta(Z)$ denotes the set of all probability measures on the finite set Z.

2. **Population games.** As already mentioned in the introduction, for the rest of this paper we will concentrate on the analysis of stochastic evolutionary game dynamics in a single unit mass population. Hence, we assume that $\bar{p} = 1$, and all agents share the same vector payoff function $F : \mathsf{X} \to \mathbb{R}^n$. We thus describe the population's aggregate behavior by a *population state* x, an element of the unit simplex $\mathsf{X} = \{x \in \mathbb{R}^n_+ | \sum_{i=1}^n x_i = 1\}$, or more specifically, the grid $\mathcal{X}^N = \mathsf{X} \cap \frac{1}{N}\mathbb{Z}^n$. The standard basis vector $e_i \in \mathbb{R}^n$ represents the *pure population state* at which all agents play strategy $i \in S = \{1, 2, \dots, n\}$.

To allow for finite-population effects, we identify a *finite-population game* with its payoff function $F^N\colon \mathcal{X}^N \to \mathbb{R}^n$, where $F^N_i(x) \in \mathbb{R}$ is the payoff to strategy i when the population state is $x \in \mathcal{X}^N$. Only the values that the function F^N_i takes on the set

$$\mathcal{X}^N_i = \{x \in \mathcal{X}^N | x_i > 0\}$$

are meaningful, since at the remaining states in $\mathcal{X}^N$ strategy i is unplayed. The *support* of a population state $x \in \mathsf{X}$ is denoted by $S_x = \{i \in S | x_i > 0\}$.

In a finite-population game, an agent who switches from strategy $i \in S_x$ to strategy $j \in S$ when the state is x changes the state to the adjacent state $x' = x + \frac{1}{N}(e_j - e_i)$. A *clever agent* should take this change in the population state into account when he considers revising his currently used strategy. To account for this, we use the *clever payoff function* $F^N_{i\to\cdot}\colon \mathcal{X}^N_i \to \mathbb{R}^n$ to denote the payoff opportunities faced by i players at each state $x \in \mathcal{X}^N_i$. The j-th component of the vector $F^N_{i\to\cdot}(x)$ is thus

$$F^N_{i\to j}(x) = F^N_j\left(x + \frac{1}{N}(e_j - e_i)\right). \tag{2}$$

Clever payoffs allow one to describe Nash equilibria of finite-population games in a simple way. The pure best response correspondence for strategy $i \in S$ in finite-population game F^N is denoted by $b_i^N : \mathcal{X}_i^N \rightrightarrows S$, and is defined by

$$b_i^N(x) = \operatorname*{argmax}_{j \in S} F_{i \to j}^N(x).$$

State $x \in \mathcal{X}^N$ is a *Nash equilibrium* of F^N if no agent can obtain a higher payoff by switching strategies: that is, $i \in b_i^N(x)$ whenever $x_i > 0$. Similarly, pure state e_i is a *strict equilibrium* if $b_i^N(e_i) = \{i\}$.

2.1. **Limits of finite-population games.** To consider large population limits, we must specify a notion of convergence for sequences $\{F^N\}_{N=N_0}^\infty$ of finite-population games. If such a sequence converges, its limit is a (*continuous*) *population game*, $F : \mathsf{X} \to \mathbb{R}^n$, which we take to be a continuous function, and hence a uniformly continuous function, from the compact set X to $\mathbb{R}$. The notion of convergence we employ for the sequence $\{F^N\}_{N=N_0}^\infty$ is uniform convergence, which asks that

$$\lim_{N\to\infty} \max_{x \in \mathcal{X}_i^N} \left\| F_{i \to \cdot}^N(x) - F(x) \right\| = 0 \qquad \forall i \in S. \tag{3}$$

In the sequel, we assume that the sequence of finite population games $\{F^N\}_{N=N_0}^\infty$ features this kind of convergence. The examples below illustrate that this mild requirement is typically satisfied in applications.

Example 1 (Random matching in normal form games)**.** The classical example for a population game is obtained from random matching of N-players playing a two-player normal form game with payoff matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$. Uniform convergence obtains in this setting, whether this occurs with self-matching ($F_i^N(x) = \sum_{j\in S} A_{ij}x_j = (\mathbf{A}x)_i$) or without ($F_i^N(x) = \frac{1}{N-1}(\mathbf{A}(Nx - e_i))_i = (\mathbf{A}x)_i + \frac{1}{N-1}((\mathbf{A}x)_i - A_{ii})$).

The pure and mixed best response correspondences for the population game F are denoted by $b : \mathsf{X} \rightrightarrows S$ and $B : \mathsf{X} \rightrightarrows \mathsf{X}$, and are defined by

$$b(x) = \operatorname*{argmax}_{i \in S} F_i(x) \quad \text{and} \quad B(x) = \operatorname*{argmax}_{y \in \mathsf{X}} y'F(x).$$

State x is a *Nash equilibrium* of F if $i \in b(x)$ whenever $x_i > 0$, or, equivalently, if $x \in B(x)$.

It follows from definition (2), the uniform convergence in (3), and the uniform continuity of F that for all strategies $i, j \in S$,

$$\begin{aligned}\lim_{N\to\infty} \max_{x \in \mathcal{X}_i^N} \left| F_{i\to j}^N(x) - F_j(x) \right| \le \lim_{N\to\infty} \max_{x \in \mathcal{X}_i^N} \Big(&\left| F_j^N(x + \tfrac{1}{N}(e_j - e_i)) \right. \\ &\left. - F_j(x + \tfrac{1}{N}(e_j - e_i)) \right| \\ &+ \left| F_j(x + \tfrac{1}{N}(e_j - e_i)) - F_j(x) \right| \Big) \\ = 0.&\end{aligned} \tag{4}$$

Thus each clever payoff function $F_{i\to\cdot}^N$ converges uniformly to F as N grows large.

Example 2 (Congestion Games)**.** An example of a population game with non-linear payoffs are congestion games. To define a congestion game [3, 40], one specifies a collection of facilities Λ (e.g. links in a highway network) and associates with each facility $\lambda \in \Lambda$ a function ℓ_λ^N describing the costs of using the facility as a function of the fraction of the population that uses it. Each action $i \in S$ (a path through the network) requires the facilities in a given subset $\Lambda_i \subseteq \Lambda$ (e.g. links on the path), and the payoff to action i is the negative of the sum of the costs accruing from these facilities. Payoffs in the resulting population game are given by $F_i^N(x) = -\sum_{\lambda\in\Lambda_i} \ell_\lambda^N(u_\lambda(x))$, where $u_\lambda(x) = \sum_{i:\lambda\in\Lambda_i} x_i$ denotes the utilization of facility λ at state x.

3. **Stochastic evolutionary game dynamics.** The evolution of the population state is described by a family of discrete-time Markov chains $\mathbf{X}^{N,\eta} = \{X_k^{N,\eta}\}_{k=0}^\infty$, whose sample paths can be described recursively as

$$\begin{aligned} &X_{k+1}^{N,\eta} = X_k^{N,\eta} + \tfrac{1}{N}\zeta_{k+1}^{N,\eta}(X_k^{N,\eta}), \quad X_0^{N,\eta} \in \mathcal{X}^N, \\ &\mathbb{P}\left(\zeta_{k+1}^{\eta}(X_k^{N,\eta}) = z | X_k^{N,\eta} = x\right) = \nu^{N,\eta}(z|x). \end{aligned}$$

The sequence of vector fields $\{\zeta_k^\eta(x)\}_{k\in\mathbb{N}}$ is i.i.d with law $\nu^{N,\eta}(\cdot|x)$. Each $\nu^{N,\eta}(\cdot|x)$ defines a probability distribution over the set of *displacement vectors* $\mathcal{Z} = \{e_j - e_i|\ i,j \in S\}$. The Markov chain $\{X_k^{N,\eta}\}_{k=0}^\infty$ takes values in the discrete grid $\mathcal{X}^N$, so that not all displacement vectors $z \in \mathcal{Z}$ are feasible at boundary points of the state space. Therefore, we introduce the set of feasible displacement vectors at state $x \in \mathsf{X}$ as $\mathcal{Z}(x) = \{z \in \mathcal{Z}| x + \varepsilon z \in \mathsf{X}$ for some $\varepsilon > 0\}$. The laws $\nu^{N,\eta}$ have the specific form

$$\nu^{N,\eta}(z|x) = \begin{cases} x_i\, \sigma_{ij}^{N,\eta}(x) & \text{if } z \in \mathcal{Z}(x) \text{ and } z = e_j - e_i, j \neq i, \\ \sum_{i=1}^n x_i\, \sigma_{ii}^{N,\eta}(x) & \text{if } z = \mathbf{0}, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

where $\sigma^{N,\eta}(x) = [\sigma_{ij}^{N,\eta}(x)]_{(i,j)\in S\times S}$ are the *conditional switch probabilities.* The conditional switch probabilities describe the information available to revising agents, and the decision rule employed for selecting a pure action. Sandholm developed a unified approach coined as *revision protocols* (see [45] and [43] for an encyclopedic overview).

More precisely, the process $\mathbf{X}^{N,\eta}$ is a Markov chain with initial condition $X_0^{N,\eta} \in \mathcal{X}^N$, transition law

$$\mathbb{P}\left(X_{k+1}^{N,\eta} = y \,\middle|\, X_k^{N,\eta} = x\right) = \sum_{z\in\mathcal{Z}} \nu^{N,\eta}(z|x)\delta_z(N(y-x)) \tag{6}$$

and generator

$$\mathcal{G}^{N,\eta} f(x) = \sum_{z\in\mathcal{Z}} [f(x + \frac{1}{N}z) - f(x)]\nu^{N,\eta}(z|x). \tag{7}$$

3.1. **Revision protocols.** In our model of evolution, each agent occasionally receives opportunities to switch actions. At such moments, an agent decides which action to play next by employing a *protocol* $\rho^\eta\colon \mathbb{R}^n \times \mathsf{X} \to \mathsf{X}^{\otimes n}$, with the choice probabilities of a current action i player being described by $\rho_{i\cdot}^\eta\colon \mathbb{R}^n \times \mathsf{X} \to \mathsf{X}$.

Specifically, if a revising action i player faces payoff vector $\pi \in \mathbb{R}^n$ at population state $x \in \mathsf{X}$, then the probability that he proceeds by playing action j is $\rho_{ij}^\eta(\pi, x)$.

Example 3 (The logit protocol)**.** A fundamental example of a revision protocol with positive choice probabilities is the *logit protocol*, defined by

$$\rho_{ij}^\eta(\pi, x) = \frac{\exp(\eta^{-1}\pi_j)}{\sum_{k\in S}\exp(\eta^{-1}\pi_k)} \tag{8}$$

for some *noise level* $\eta > 0$. When η is small, an agent using this protocol is very likely to choose an optimal action, but places positive probability on every action, with lower probabilities being placed on worse-performing actions.

Example 4 (Perturbed best response protocols)**.** One can generalize (8) by assuming that agents choice probabilities maximize the difference between their expected base payoff and a convex penalty:

$$\rho^{\eta}_{i\cdot}(\pi, x) = \operatorname*{argmax}_{x \in \mathsf{X}} \left(\sum_{k \in S} \pi_k x_k - \eta h(x) \right),$$

where $h\colon \mathsf{X} \to [-\infty, \infty]$ is a proper lower semi-continuous convex function satisfying the typical properties of a Mirror Descent algorithm (see [34]). Specifically, we require that h is continuously differentiable on the relative interior of X, and that $\|\nabla h(x)\|$ approaches infinity whenever x approaches the boundary of X. The logit protocol (8) is recovered when h is the negated entropy function $\sum_{k \in S} x_k \ln(x_k)$.

Example 5 (The pairwise logit protocol)**.** Under the *pairwise logit protocol*, a revising agent chooses a candidate action at random, and then applies the logit rule (8) only to his current action and the candidate action:

$$\rho^{\eta}_{ij}(\pi, x) = \frac{\exp(\eta^{-1}\pi_j)}{\exp(\eta^{-1}\pi_i) + \exp(\eta^{-1}\pi_j)}$$

Example 6 (Imitation with "mutations")**.** Suppose that with probability $1 - \varepsilon$, a revising agent picks an opponent at random and switches to her action with probability proportional to the opponent's payoff, and that with probability $\varepsilon > 0$ the agent chooses an action at random. If payoffs are normalized to take values between 0 and 1, the resulting protocol takes the form

$$\rho^{\eta}_{ij}(\pi, x) = \begin{cases} (1-\varepsilon)\,\frac{N}{N-1} x_j\, \pi_j + \frac{\varepsilon}{n} & \text{if } j \neq i \\ (1-\varepsilon)\left(\frac{N x_i - 1}{N} + \sum_{k \neq i} \frac{N}{N-1} x_k (1 - \pi_k)\right) + \frac{\varepsilon}{n} & \text{if } j = i. \end{cases}$$

The positive mutation rate ensures that all actions are chosen with positive probability.

Example 7 (Direct Exponential Protocols)**.** An all encompassing model of stochastic evolution is provided by revision protocols defined in terms of exponential functions

$$\rho^{\eta}_{ij}(\pi, x) = \frac{\exp(\eta^{-1}\psi(\pi_i, \pi_j))}{\kappa_{ij}(\pi)}, \tag{9}$$

where $\psi : \mathbb{R}^2 \to \mathbb{R}$ is a given function satisfying

$$\begin{aligned} &\psi(\pi_i, \pi_j) - \psi(\pi_j, \pi_i) = \pi_j - \pi_i, \text{ and} \\ &\kappa_{ij}(\pi) = \kappa_{ji}(\pi). \end{aligned}$$

Examples for direct exponential protocols are directly given by the following specifications

- Logit: $\psi(\pi_i, \pi_j) = \pi_j$;
- Pairwise Comparison: $\psi(\pi_i, \pi_j) = \frac{1}{2}(\pi_j - \pi_i)$;
- Simulated Annealing: $\psi(\pi_i, \pi_j) = [\pi_i - \pi_j]_+$.

For many further examples of revision protocols, see [43]. Whatever specification of a revision protocol one adopts, the conditional switch probabilities are generated via the identity

$$\sigma_{ij}^{N,\eta}(x) = \rho_{ij}^{\eta}(F_{i\to\cdot}^{N}(x), x). \tag{10}$$

3.2. **Scaling assumptions.** Our large deviations results concern the behavior of sequences $\mathbf{X}^{N,\eta}$ of Markov chains as the scaling parameters (N, η) approach their extreme values. We assume that there is a Lipschitz continuous function $\sigma^{N,\eta}\colon \mathcal{X}^N \to \mathbb{R}_+^{n\times n}$ that describes the limiting switch probabilities, in the sense that

$$\lim_{N\to\infty} \max_{x\in\mathcal{X}^N} \max_{i,j\in S} |\sigma_{ij}^{N,\eta}(x) - \sigma_{ij}^{\eta}(x)| = 0. \tag{11}$$

In addition, we assume that limiting switch probabilities are bounded away from zero: there is a $\varsigma > 0$ such that

$$\min_{x\in\mathsf{X}} \min_{i,j\in S} \sigma_{ij}^{\eta}(x) \geq \varsigma. \tag{12}$$

This is so under all of the revision protocols from Section 3.1. Assumption (12) and the transition law (5) imply that the Markov chain $\mathbf{X}^{N,\eta}$ is aperiodic and irreducible, for N large enough. Thus for such N, $\mathbf{X}^{N,\eta}$ admits a unique stationary distribution, $\mu^{N,\eta}$, which is both the limiting distribution of the Markov chain and its limiting empirical distribution along almost every sample path. This assumption rules out frequency dependent selection via the Moran process, the dominant mathematical model in theoretical biology [54, 56], in which absorbing boundary states are present. Small perturbations of the base process, however, would satisfy Assumption (12).

Assumptions (11) and (12) imply that the transition kernels (5) of the Markov chains $\mathbf{X}^{N,\eta}$ approach a limiting kernel $\nu^{\eta}\colon \mathsf{X} \to \mathcal{P}(\mathcal{Z})$, defined by

$$\nu^{\eta}(z|x) = \begin{cases} x_i\,\sigma_{ij}^{\eta}(x) & \text{if } z = e_j - e_i \text{ and } j \neq i \\ \sum_{i\in S} x_i\,\sigma_{ii}^{\eta}(x) & \text{if } z = \mathbf{0}, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Condition (11) implies that the convergence of $\nu^{N,\eta}$ to ν^{η} is uniform:

$$\lim_{N\to\infty} \max_{x\in\mathcal{X}^N} \max_{z\in\mathcal{Z}} \left|\nu^{N,\eta}(z|x) - \nu^{\eta}(z|x)\right| = 0. \tag{14}$$

The probability measures $\nu^{\eta}(\cdot|x)$ depend Lipschitz continuously on x, and by virtue of condition (12), each measure $\nu^{\eta}(\,\cdot\,|x)$ has support $\mathcal{Z}(x)$.

4. **The small noise double limit.** Traditionally, stochastic evolutionary game dynamics have been analyzed in regimes where the noise level η is small and the population size N is large. This limiting scenario is tractable essentially because the parameter η controls the randomness of the Markovian dynamics in all known applications. In analogy with spin glass models, η is the temperature parameter of the system, the limit $\eta \to 0$ is a cooling phase of the dynamics. At least formally, the limit $\eta \to 0$ corresponds to deterministic dynamics, and the possible transitions are easily identifiable. The limit $N \to \infty$ yields formally tractable results by exploiting the classical principle of least action: To learn which transitions are the ones the limiting dynamics will take, we have to find a curve of minimal resistance. To make this precise, the papers [52] and [46] developed a rigorous optimal control framework to identify these paths of minimal resistance. However, the optimal control problems are challenging as the control variables enter the problem in a linear way, and hence bang-bang solutions or singular arcs are likely to appear. In fact [52] identifies

a singular arc as the optimal solution in the logit dynamics. More generally, the recent paper [48] investigates the arising optimal control problem with the help of Hamilton-Jacobi equations and provides a characterization of solutions in terms of obstacle problems.

4.1. **Analysis of the small noise double limit.** For concreteness, we focus in this section on the special case where agents are matched against all opponents to play a symmetric two-player normal form game $\mathbf{A} \in \mathbb{R}^{n\times n}$, where A_{ij} is the payoff that an agent playing i obtains when matched against an agent playing j. During such a matching, the payoff obtained by a current action i player is

$$F_i(x) = \sum_j A_{ij}x_j = (\mathbf{A}x)_i.$$

State $x^* \in \mathsf{X}$ is a *Nash equilibrium* of $\mathbf{A}$ if all strategies in use at x^* are optimal:

$$i \in \operatorname*{argmax}_{j\in S} (\mathbf{A}x^*)_j \text{ whenever } x_i^* > 0. \tag{15}$$

We can characterize Nash equilibrium in terms of the *best response regions*

$$\mathcal{B}^i = \{x \in \mathsf{X} \colon (\mathbf{A}x)_i \geq (\mathbf{A}x)_j \text{ for all } j \in S\} \tag{16}$$

in which each action i is optimal. Specifically, x^* is a Nash equilibrium if $x^* \in \mathcal{B}^i$ whenever $x_i^* > 0$.

The population process $\mathbf{X}^{N,\eta}$ is assumed to be generated by target protocols in the sense that the conditional switch probabilities depend only on the targeted strategy. This means, by an obvious abuse of notation, that

$$\nu^{N,\eta}(e_j - e_i|x) = x_i\rho_j^\eta(F^N_{i\to\cdot}(x)) \qquad \forall x \in \mathcal{X}^N, i, j \in S.$$

We assume that the probability of playing an action that is suboptimal at payoff vector $\pi \in \mathbb{R}^n$ vanishes at a well-defined rate as η approaches zero. These rates are captured by the *unlikelihood function* $\Upsilon \colon \mathbb{R}^n \to \mathbb{R}^n_+$, which is defined by

$$(\forall j \in S): \quad \Upsilon_j(\pi) = -\lim_{\eta\to 0} \eta \log \rho_j^\eta(\pi) \tag{17}$$

In words, $\Upsilon_j(\pi)$ is the rate of decay of the probability that action j is chosen as η approaches zero.

Example 8 (Logit choice)**.** It is easy to verify that the logit choice protocol has piecewise linear unlikelihood function given by

$$\Upsilon_j(\pi) = \max_{k\in S} \pi_k - \pi_j \tag{18}$$

Example 9 (Probit choice)**.** The *probit choice protocol* is an additive random utility model in which the payoff vector π is perturbed by adding a standard normal random vector scaled by the noise level η. The probability of choosing strategy j is given by:

$$\rho_j^\eta(\pi) = \mathbb{P}\left(j \in \operatorname*{argmax}_{k\in S}(\pi_k + \eta U_k)\right) \tag{19}$$

where $\{U_k\}_{k\in S}$ are independent standard normal random variables. Using Sanov's Theorem, [14] derives the unlikelihood function associated with this choice function as a piecewise quadratic function.

To understand the communication structure of the Markovian dynamics $\mathbf{X}^{N,\eta}$ when $\eta \to 0$, we assign costs to paths γ through the state space. For finite N and $\eta \to 0^+$, observe that the probability that the Markov chain $\mathbf{X}^{N,\eta}$ follows a path $\{x_0, x_1, \ldots, x_n\}$ is $\prod_{i=1}^n (\mathbf{P}^{N,\eta})_{x_{i-1},x_i}$, where $\mathbf{P}^{N,\eta}_{x,y} = \mathbb{P}(X_1^{N,\eta} = y | X_0^{N,\eta} = x)$ for $(x,y) \in \mathcal{X}^N \times \mathcal{X}^N$. Computing the log-likelihood of the given path gives for small η the approximation

$$-\eta \sum_{i=1}^n \ln(\mathbf{P}^{N,\eta}_{x_{i-1},x_i}) \approx \begin{cases} \sum_{i=1}^n \langle \Upsilon(F(x_i)), [N(x_i - x_{i-1})]_+ \rangle & \text{if } \prod_{i=1}^n \mathbf{P}^{N,\eta}_{x_{i-1},x_i} > 0, \\ +\infty & \text{else.} \end{cases}$$

Here $[u]_+ = ([u_1]_+, \ldots, [u_n]_+)$ is the positive part applied componentwise to a vector $u \in \mathbb{R}^n$. Motivated by the above structure of the approximate path cost function for $\eta \to 0$, [46] define the cost functional as follows:

$$c_{x,T}(\gamma) = \begin{cases} \int_0^T \langle \Upsilon(F(\gamma(t))), [\dot\gamma(t)]_+ \rangle \mathrm{d}t & \text{if } \gamma \in \mathbf{AC}_x[0,T], \\ +\infty & \text{otherwise} \end{cases} \tag{20}$$

The scalar $\Upsilon_j(F(\gamma(t)))$ measures how unlikely it is at state $\gamma(t)$ that a revising agent chooses action j. The inner product of $\Upsilon(F(\gamma(t)))$ with $[\dot\gamma(t)]_+$ is a weighted sum of these unlikelihoods, with non-negative weights given by how quickly each strategy's representation is supposed to grow on path γ at time t; strategies that become less common are ignored. The aggregate of these weighted sums over the path γ is the path's cost.

The running cost $L(x,u)$ associated with cost function given by Eq. (20) is

$$L(x,u) = \begin{cases} \langle \Upsilon(F(x)), [u]_+ \rangle & \text{if } u \in \mathsf{T}_{\mathsf{X}}(x), \\ +\infty & \text{otherwise} \end{cases} \tag{21}$$

where $\mathsf{T}_{\mathsf{X}}(x)$ is the tangent cone of X at $x \in \mathsf{X}$.

Let $\mathsf{A}, \mathsf{B} \subset \mathsf{X}$ and $T > 0$. $\Gamma_T(\mathsf{A}, \mathsf{B})$ denote the set of all continuous paths through X starting at some point in A and terminating at some point in B in finite time T. Formally, we have

$$\Gamma_T(\mathsf{A},\mathsf{B}) = \{\gamma \in \mathbf{C}[0,T] | \gamma(0) \in \mathsf{A}, \gamma(T) \in \mathsf{B} \text{ and } \gamma(t) \in \mathsf{X}\ \forall t \in [0,T]\}.$$

Let $x \in \mathsf{X}$ and $\mathsf{O} \subset \mathsf{X}$. Following [16], we introduce two quantities, *radius* and *coradius* of x with respect to the set O. The radius of the set O, denoted by $\mathsf{rad}(x, \mathsf{O})$, measures the difficulty of reaching O starting from an arbitrary state x. The coradius, $\mathsf{corad}(\mathsf{O}, x)$, measures the difficulty of leaving the source set O to reach a target state $x \in \mathsf{X}$. Formally, we have

$$\mathsf{rad}(x,\mathsf{O}) = \inf\{c_{x,T}(\gamma) | \gamma \in \Gamma_T(\{x\}, \mathsf{O}), T > 0\}, \tag{22}$$

and the coradius

$$\mathsf{corad}(\mathsf{O},x) = \inf\{c_{\gamma(0),T}(\gamma) | \gamma \in \Gamma_T(\mathsf{O}, \{x\}), T > 0\}. \tag{23}$$

If the set O is understood from the context, we can suppress them from the definitions of these two value functions. It is then easy to see that $\mathsf{rad}(x,\mathsf{O})$ corresponds to the target problem, and $\mathsf{corad}(\mathsf{O},x)$ to the source problem, respectively, studied in [48]. They provide a PDE characterization of the radius and coradius of the Lipschitz continuous subsolutions of the Hamilton-Jacobi equation

$$H(x, \nabla\phi(x)) = 0, \tag{24}$$

involving the Fenchel conjugate

$$H(x,p) = \sup_{u \in \mathsf{U}} \{\langle u, p \rangle - L(x,u)\}, \tag{25}$$

where U is a compact convex subset of $\mathsf{T}_{\mathsf{X}}(x)$. The exact statement is as follows:

Theorem 4.1 ([48]). *The radius $x \mapsto \mathsf{rad}(x, \mathsf{T})$ is the maximal Lipschitz continuous function $R : \mathsf{X} \to \mathbb{R}$ satisfying*

$$\begin{cases} H(x, -\nabla R(x)) \leq 0 & \text{for almost all } x \in \mathsf{X}^\circ, \\ R(x) = 0 & \text{for all } x \in \mathsf{T}. \end{cases} \tag{26}$$

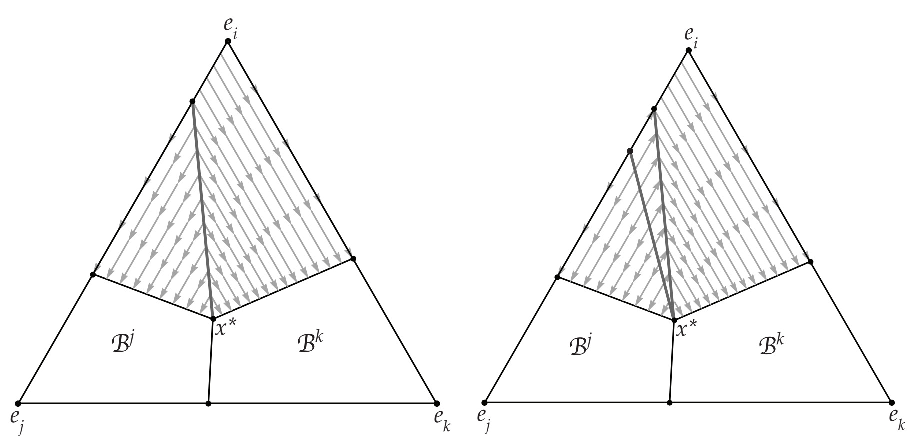
The coradius $x \mapsto \mathsf{corad}(\mathsf{S}, x)$ is the maximal Lipschitz continuous function $W : \mathsf{X} \to \mathbb{R}$ satisfying

$$\begin{cases} H(x, \nabla W(x)) \leq 0 & \text{for almost all } x \in \mathsf{X}^\circ, \\ W(x) = 0 & \text{for all } x \in \mathsf{S}. \end{cases} \tag{27}$$

where $\mathsf{X}^\circ = \{x \in \mathsf{X} | x_i > 0 \text{ for all } i = 1, 2, \ldots, n\}$.

With the help of this PDE characterization, [48] go further in deriving a powerful verification theorem which turns out to be a convenient analytical tool to solve non-trivial examples. In the next sections we illustrate this in the context of three-strategy games. However, there is more research needed in order to really complete our understanding of the model in this limiting case. In particular, the next step must be the development of efficient numerical solution techniques for quantifying the radius and coradius of an equilibrium.

Open Problem 1. *Develop an efficient numerical scheme for computing the radius and coradius of a Nash equilibrium using the Hamilton-Jacobi equations from Theorem 4.1.*



(A) Direct exit (B) Indirect exit

FIGURE 1. Optimal exit paths with non-binding state constraints.

4.2. **An exit problem in three-strategy coordination games.** In this section, we summarize the solution to the exit cost problem in a class of symmetric three-strategy coordination games under the logit choice and probit choice rules in the small noise double limit. The analysis in this section is based on [46] and [1].

We denote an element in the three-strategy coordination games under consideration by a 3×3 matrix $\mathbf{A}$. We suppose that $\mathbf{A}$ is a *coordination game*: $A_{ii} > A_{ji}$ for all $i, j \in S$, $i \neq j$. This implies that if one's match partner plays i, one is best off playing

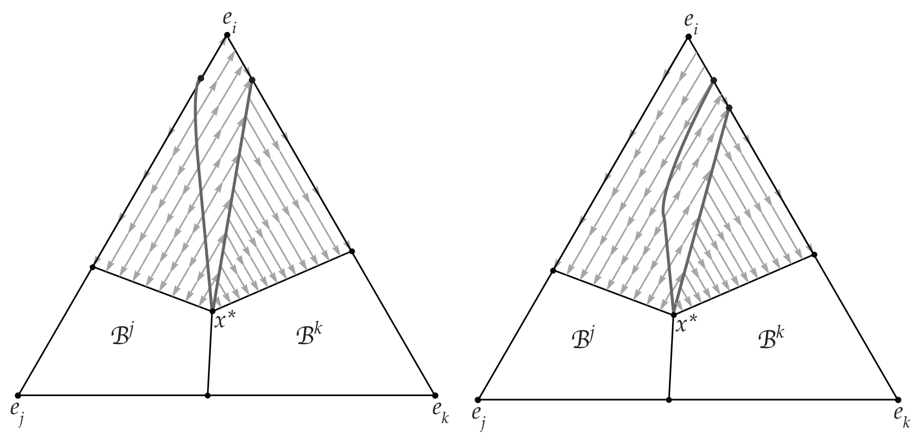


FIGURE 2. Optimal exit paths with indirect exit and binding state constraints.

i oneself. Therefore each pure strategy i will be a strict Nash equilibrium. We also suppose that $\mathbf{A}$ has *marginal bandwagon property* of [35]: $A_{ii} - A_{ik} > A_{ji} - A_{jk}$ for all $i, j, k \in S$ with $i \notin \{j, k\}$. In words, the above condition requires that when some agent switches to strategy i from any other strategy k, current strategy i players benefit most.

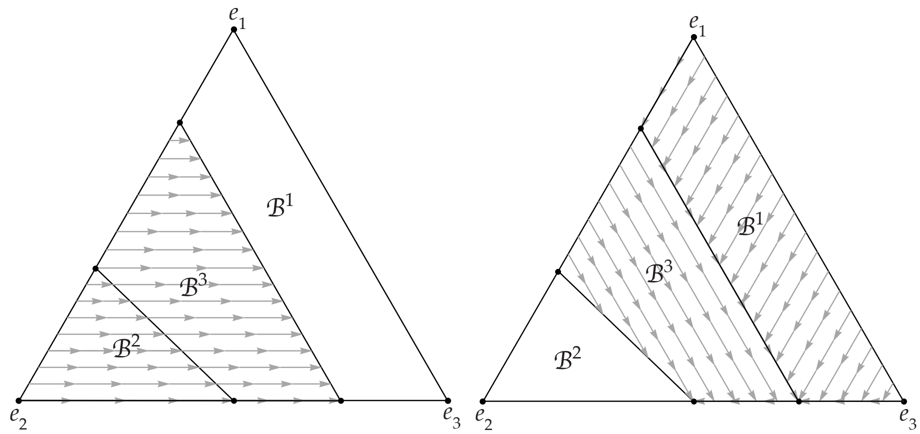
For $x \in \mathcal{B}^i$, the exit cost problem is defined by

$$C_{\text{exit}}(x) = \mathsf{rad}(x, \mathcal{B}^j \cup \mathcal{B}^k) \tag{28}$$

The solution to this exit problem can be used to assess the expected time until the evolutionary process leaves the basin of attraction of a stable equilibrium. When N is sufficiently large, the exponential growth rate of the expected waiting time to leave the initial basin of attraction $\mathcal{B}^i$ as η vanishes is approximately $N \times C_{\text{exit}}(e_i)$.

[46] showed that the continuous paths which solve the exit problem i.e., the *optimal exit paths* under logit choice rule will be as shown in Figure 1, panel (A). The best-response region $\mathcal{B}^i$ is split into two regions; in one, the optimal control is $e_j - e_i$ and exit paths lead to the boundary of best response region of strategies i and j, $\mathcal{B}^{ij}$; in the other the optimal control is $e_k - e_i$ and exit paths lead to the boundary of best response region of strategies i and k, $\mathcal{B}^{ik}$. The boundary between the regions is a ray whose endpoint is the unique interior mixed equilibrium x^*. From points on this ray, motion in either basic direction $e_j - e_i$ or $e_k - e_i$ is optimal. The optimal exit paths under logit choice will always feature *direct exit* i.e., from any initial condition in the basin of attraction, the optimal exit paths will involve the agents always switching away from the status quo equilibrium strategy.

Using the verification theorem of [48], [1] showed that the *optimal exit paths* under probit choice may lead to new qualitative features. In particular, [1] showed that these paths may feature *indirect exit* i.e., the optimal exit paths from some initial conditions in the basin of attraction may involve the agents switching back to the status quo equilibrium strategy as shown in Figure 1, panel (B) and Figure 2. These cases can be further subdivided into cases with non-binding state constraints as shown in Figure 1, panel (B) and binding state constraints as shown in Figure 2.



(A) Transitions to state 1

(B) Transitions to state 2

FIGURE 3. Optimal transition paths in Bilingual games.

4.3. **A transition problem in bilingual games.** In this section, we summarize the solution to the transition cost problem in a class of three-strategy games called as bilingual games under the logit choice rule in the small noise double limit. The analysis in this section is based on [2].

Consider a two-strategy coordination game with payoffs as shown in Table 1. Assume that $c \leq d < a < c + d$. Strategy 1 is then the payoff dominant equilibrium, while strategy 2 is the risk dominant equilibrium.[1] Existing results show that the risk dominant equilibrium (inefficient equilibrium, strategy 2) is selected in the long run under the perturbed best response protocols (see [7]) which includes the logit choice model. This is because in the above class of games, it is relatively easier to breakdown the efficient equilibrium compared to breaking down the inefficient equilibrium as the cost of transition from the risk dominant equilibrium to the payoff dominant equilibrium is greater than the cost of transition from payoff dominant equilibrium to the risk dominant equilibrium when agents employ noisy best response models like the logit choice rule.

	1	2
1	a, a	$0, c$
2	$c, 0$	d, d

TABLE 1. Coordination Game

If we allow an extra option of adopting both strategies at an additional cost $\epsilon > 0$ in the game from Table 1, then we have a new game whose payoffs will be as shown in Table 2. These games are known as *bilingual games* [37], and strategy 3 is called as the *bilingual option*. For low adoption costs ϵ, the bilingual option can be thought of as a hedging strategy in the face of uncertainty about what other

[1]If a player has a uniform prior over other players' strategies, then the risk dominant strategy is the one which maximizes his expected payoff [19].

players choose. Although the bilingual option, strategy 3 is not a Nash equilibrium, its presence can change the equilibrium which is selected in the long run by altering the transition costs between the strict equilibria.

	1	2	3
1	a, a	$0, c$	$a, a-\epsilon$
2	$c, 0$	d, d	$d, d-\epsilon$
3	$a-\epsilon, a$	$d-\epsilon, d$	$a-\epsilon, a-\epsilon$

TABLE 2. Bilingual Game

Let $i, j \in \{1,2\}, i \neq j$. For $x \in \mathcal{B}^j$, the transition cost problem in a bilingual game to the strict equilibrium e_i is defined by:

$$C_{\text{tran}}(x, e_i) = \mathsf{rad}(x, \mathcal{B}^i) \tag{29}$$

The solution to the transition problem (Eq. (29)) gives the stochastically stable state i.e., the equilibrium selected in the long run. Strict equilibrium e_1 will be stochastically stable if and only if $C_{\text{tran}}(e_1, e_2) \geq C_{\text{tran}}(e_2, e_1)$. Similarly, strict equilibrium e_2 will be stochastically stable if and only if $C_{\text{tran}}(e_2, e_1) \geq C_{\text{tran}}(e_1, e_2)$.

Using the verification theorem of [48], [2] showed that in a bilingual game when the adoption cost ϵ is sufficiently low, the continuous paths which solve the transition problem i.e., the *optimal transition paths* under logit choice will be as shown in Figure 3. By explicitly computing the transition costs between the strict equilibria e_1 and e_2, [2] showed that if the adoption cost of the bilingual option ϵ is sufficiently low, then the payoff dominant equilibrium e_1 will be selected in the long run under the logit choice rule in the small noise double limit.

5. Large population analysis.

5.1. Mean dynamics.

The evolutionary game dynamics $\mathbf{X}^{N,\eta}$ was defined in discrete time and to take values in the discrete set of finite population states $\mathcal{X}^N$. We interpret this process as a data-generating process. Real time of play is measured on the set $\mathbf{T}^N = \{0, 1/N, 2/N, \ldots,\}$, and at each point $t_k^N = k/N \in \mathbf{T}^N$ a single player is granted a revision opportunity. Consider the continuous-time process

$$\hat{X}^{N,\eta}(t) = X_k^{N,\eta} + (Nt-k)(X_{k+1}^{N,\eta} - X_k^{N,\eta}) \qquad \forall t \in [t_k^N, t_{k+1}^N], k \in \mathbb{N}_0. \tag{30}$$

The process $\{\hat{X}^{N,\eta}(t); t \geq 0\}$ is a random element of the set $\mathbf{C}(\mathbb{R}_+ : \mathsf{X})$. Define the mean increment of the stochastic process $\mathbf{X}^{N,\eta}$ by

$$V^{N,\eta}(x) = \sum_{z \in \mathcal{Z}} z \nu^{N,\eta}(z|x) = \sum_{i,j} (e_j - e_i) x_i \sigma_{ij}^{N,\eta}(x), \tag{31}$$

so that for each $j \in S$ we have

$$V_j^{N,\eta}(x) = \sum_{i=1}^{n} x_i \sigma_{ij}^{N,\eta}(x) - x_j.$$

For each population size N, the vector field $V^{N,\eta}$ is defined only on the discrete set $\mathcal{X}^N$. However, as N becomes large the sequence $\{V^{N,\eta}\}_{N \geq N_0}$ approaches the function

$$V^{\eta}(x) = \sum_{z \in \mathcal{Z}} z \nu^{\eta}(z|x),$$

where the family of probability distributions $\{\nu^\eta(\cdot|x) : x \in \mathsf{X}\}$ is defined by (13).[2]

The function $x \mapsto V^\eta(x) \in \mathsf{T}_{\mathsf{X}}(x)$ is bounded and continuous on X. The *mean-field dynamic* associated with the evolutionary game process $\mathbf{X}^{N,\eta}$ is defined as the unique solution of the initial value problem

$$\dot{\mathbf{x}}(t) = V^\eta(\mathbf{x}(t)),\ \mathbf{x}(0) = x \in \mathsf{X}. \tag{M}$$

This mean dynamic corresponds to the perturbed best-response dynamics studied in [5] and [24]. The induced semi-flow is defined by a mapping $\phi : \mathbb{R}_+ \times \mathsf{X} \to \mathsf{X}$ satisfying

$$\begin{aligned} &\phi(0, x) = x \qquad \forall x \in \mathsf{X}, \text{ and} \\ &\frac{\mathrm{d}\phi(t, x)}{\mathrm{d}t} = V^\eta(\phi(t, x)) \qquad \forall t \geq 0. \end{aligned}$$

Let us give some examples for mean-dynamics, which have been studied in the literature.

Example 10 (The Logit dynamics)**.** The Logit dynamics is the most common perturbation of the discontinuous best-response differential inclusion. It is defined by the vector field

$$V_i^\eta(x) = \frac{\exp\left(\frac{1}{\eta} F_i(x)\right)}{\sum_{j \in S} \exp\left(\frac{1}{\eta} F_j(x)\right)} - x_i \quad \forall i \in S \tag{32}$$

The central role played by the mean-field dynamics is given by the following finite-horizon approximation result due to [5, 6].

Theorem 5.1 ([5])**.** *There exists a constant $c > 0$ such that for all $\varepsilon > 0, T < \infty, x^N \in \mathcal{X}^N$, and N large enough, we have that*

$$\mathbb{P}_{x^N}\left(\max_{t \in [0,T] \cap \mathbf{T}^N} \left\| \hat{X}^{N,\eta}(t) - \phi(t, x) \right\|\right) \leq 2n \exp(-\varepsilon^2 c N).$$

A *rest-point* for the mean-field dynamics (M) is a population state $x^* \in \mathsf{X}$ such that $V^\eta(x^*) = \mathbf{0}$. A rest point x^* is said to be an attractor if there exists a full neighborhood $B(x^*; \varepsilon)$ (in X) from which all solution trajectories to (M) converge to x^*. The *basin of attraction* of a rest point x^* is defined as

$$\mathcal{B}(x^*) = \{x \in \mathsf{X} | \lim_{t \to \infty} \|\phi(t, x) - x^*\| = 0\}.$$

Since [5] it is known that attractors of the mean dynamics (M) are good predictors for the pattern of play in large population settings and over finite (but possibly very large) time lengths. In particular, they show that once the deterministic flow enters a small neighborhood contained in the basin of attraction of an attracting rest point x^*, the probability that the stochastic process escapes this neighborhood approaches 0 as the population size becomes large. Conversely, if a rest point is not attractive, then the stochastic process escapes an arbitrarily small neighborhood of this rest point in finite time with probability 1.

[2]That this is indeed the case follows immediately from the continuity properties of the noisy-best response function σ^η, and the uniform convergence of the sequence of finite population games.

5.2. **Large deviation principle.** In [47], a sample-path large deviations principle has been derived for the sequence of Markovian population processes $\mathbf{X}^{N,\eta}$. To state the large deviation principle for the sequence of interpolated processes, we must introduce a rate function and some general definitions and terminology.

Definition 5.2. A curve $\gamma : [0,T] \to \mathsf{X}$ is absolutely continuous if there exists a function $u \in L^1([0,T] : \mathbb{R}^n)$ such that $\gamma(t) = \gamma(0) + \int_0^t u(s)\mathrm{d}s$. We write $\dot{\gamma}(t) = u(t)$, understanding that this is interpreted in an a.e. sense. A function $\gamma : [0,\infty) \to \mathsf{X}$ is absolutely continuous if it is absolutely continuous for every $T \geq 0$.

Define the log-moment generating function as

$$H^\eta(x,p) = \log\left(\sum_{z\in\mathcal{Z}} e^{\langle p,z\rangle}\nu^\eta(z|x)\right). \tag{33}$$

It is well known that $H^\eta(x,\cdot) : \mathbb{R}^n_0 \to \mathbb{R}$ is a continuous convex function. Using convex duality arguments, we can define its conjugate function as

$$L^\eta(x,u) = \sup_{p\in\mathbb{R}^n_0}\{\langle p,u\rangle - H^\eta(x,p)\}. \tag{34}$$

This function is crucial for deriving sample path large deviation estimates. The following Lemma summarizes its main properties, and establishes also a connection with the relative entropy function

$$R(q||\nu) = \sum_{z\in\mathcal{Z}} q(z)\log\frac{q(z)}{\nu(z)}$$

for $q,\nu \in \Delta(\mathcal{Z})$. In the following, we denote by Z the convex hull generated by the finite set $\mathcal{Z}$, and analogously $\mathsf{Z}(x) = \mathrm{conv}(\mathcal{Z}(x))$.

Lemma 5.3. *The function $L^\eta : \mathsf{X}\times\mathsf{Z} \to [0,\infty]$ has the following properties.*

(i) Its effective domain is the set $\mathrm{dom}(L^\eta) = \{(x,u) \in \mathsf{X}\times\mathsf{Z} | u \in \mathsf{Z}(x)\}$, and for every $x \in \mathsf{X}$, $\mathrm{dom}(L^\eta(x,\cdot)) = \mathsf{Z}(x)$;

(ii) $L^\eta(x,\cdot)$ is non-negative and convex. Additionally, it is lower semi-continuous in $(x,u) \in \mathsf{X}\times\mathsf{Z}$;

(iii) It has the representation

$$L^\eta(x,u) = \inf\{R(q||\nu^\eta(\cdot|x))|q \in \Delta(\mathcal{Z}), \sum_{z\in\mathcal{Z}} zq(z) = u\}. \tag{35}$$

In particular, $L^\eta(x,V^\eta(x)) = 0$ for all $x \in X$.

Proof. (i) This follows from (iii).

(ii) For fixed $x \in \mathsf{X}$, let $H^\eta_x(p) \equiv H^\eta(x,p)$. An application of Hölder's inequality shows that for every $p,p' \in \mathbb{R}^n_0$ and $\lambda \in (0,1)$

$$\sum_{z\in\mathcal{Z}}\nu^\eta(z|x)\exp(\langle p,z\rangle)^\lambda\exp(\langle p',z\rangle)^{1-\lambda}$$
$$\leq \left(\sum_{z\in\mathcal{Z}}\nu^\eta(z|x)\exp(\langle p,z\rangle)\right)^\lambda\left(\sum_{z\in\mathcal{Z}}\nu^\eta(z|x)\exp(\langle p',z\rangle)\right)^{1-\lambda}$$

Hence,

$$H^\eta_x(\lambda p + (1-\lambda)p') \leq \lambda H^\eta_x(p) + (1-\lambda)H^\eta_x(p').$$

Thus, $H^\eta_x(p)$ is convex in p. The convexity of $L^\eta_x(\cdot) \equiv L^\eta(x,\cdot)$ follows then from Theorem 12.2 in [39]. The dual formula of Fenchel-Legendre transform [39], chapter 12, gives

$$H^\eta(x,p) = \sup_{u\in Z}\{\langle p,u\rangle - L^\eta(x,u)\}.$$

Since $H^\eta(x,0) = \log\sum_{z\in\mathcal{Z}}\nu^\eta(z|x) = 0$, it follows that $L^\eta(x,u)\geq 0$, and $\inf_{u\in\mathsf{Z}} L^\eta(x,u) = 0$. The function L^η is continuous on the relative interior of its domain. It remains to check the continuity property at the relative boundary of $\mathrm{dom}(L^\eta)$. Fix a point $x\in\mathrm{bd}(\mathsf{X})$. If $u\in\mathsf{Z}(x)$, then there exists a sequence $x^n\to x$ such that $L^\eta(x^n,u)<\infty$ for all $n\in\mathbb{N}$. If $u\notin\mathsf{Z}(x)$, then since $\mathsf{Z}(x)$ is a compact convex set, there exists a $\varepsilon>0$ such that $\mathrm{dist}(u,\mathsf{Z}(x))\geq\varepsilon>0$. By the separating hyperplane theorem, we can find a full neighborhood around u, relative to Z which is disjoint from $\mathsf{Z}(x)$. Hence,

$$\begin{aligned}\liminf_{(x',u')\to(x,u)} L^\eta(x',u') &= \lim_{\varepsilon\to 0^+}\left[\inf\{L^\eta(x',u')|\,\|(x',u')-(x,u)\|<\varepsilon\}\right]\\ &= \infty = L^\eta(x,u).\end{aligned}$$

(iii) The first statement is part (f) of Lemma 6.2.3 in [15]. For the second statement observe that $V^\eta(x) = \sum_z z\nu^\eta(z|x)$. Hence

$$\inf\{R(q||\nu^\eta(\cdot|x))|q\in\Delta(\mathcal{Z}), \sum_{z\in\mathcal{Z}} zq(z) = V^\eta(x)\} = 0,$$

by choosing $q(z) = \nu^\eta(z|x)$.

□

We define the *path cost function* $c^\eta_{x,T}\colon \mathbf{C}[0,T]\to[0,\infty]$ by

$$c^\eta_{x,T}(\gamma) = \begin{cases}\int_0^T L^\eta(\gamma(t),\dot\gamma(t))\,\mathrm{d}t & \text{if } \gamma\in\mathbf{AC}_x[0,T],\\ \infty & \text{otherwise.}\end{cases} \tag{36}$$

Following the general theory of [15], we conjecture that $c^\eta_{x,T}$ serves as a rate function for the sample path large deviations principle. For that we have to verify some a-priori stability properties of this functional.

Proposition 1. *Given (η,x,T), the function $c^\eta_{x,T}:\mathbf{C}([0,T]:\mathsf{X})\to[0,\infty]$ has compact level sets, i.e. the sets of the form*

$$\Gamma(M) = \{\gamma\in\mathbf{C}([0,T]:\mathsf{X})|c^\eta_{x,T}(\gamma)\leq M\}$$

for $M\in\mathbb{R}_+$ are compact in sup-norm topology.

Proof. Let $\{\gamma^n\}_{n=1}^\infty$ be a sequence of functions with $\gamma^n\in\Gamma(M)$. Then $c^\eta_{x,T}(\gamma^n)<\infty$ for all n. Hence, $\{\gamma^n\}$ is a sequence of absolutely continuous functions with $\dot\gamma^n(t)\in\mathsf{Z}(\gamma(t))$ a.e. $t\in[0,T]$. The Arzelá-Ascoli theorem gives a uniformly convergent subsequence $\{\gamma^{n_j}\}$ with limit $\gamma\in\mathbf{C}([0,T]:\mathsf{X})$. For every n, there exists a measurable function $q^n:[0,T]\to\Delta(\mathcal{Z})$ such that $\dot\gamma^n(t) = \sum_{z\in\mathcal{Z}} zq^n_t(z)$ holds for every $t\in[0,T]$. Consider the space

$$\mathcal{A}_T := \{\alpha:[0,T]\to\Delta(\mathcal{Z})|\alpha(\cdot)\text{ measurable}\}.$$

By Alaoglu's Theorem, this set is weak*-compact. Hence $\{q^n\} \subset \mathcal{A}_T$ has a weak* convergent subsequence $\{q^{n_j}\}_{j=1}^{\infty}$, meaning that

$$\lim_{n\to\infty} \int_0^T \sum_{z\in\mathcal{Z}} g(t,z) q_t^{n_j}(z) \mathrm{d}t = \int_0^T \sum_{z\in\mathcal{Z}} g(t,z) q_t(z) \mathrm{d}t \tag{37}$$

for every bounded continuous function $g : [0,T] \times \mathcal{Z} \to \mathbb{R}$. Fix such a convergent subsequence, and set

$$\gamma^{n_j}(t) = x + \int_0^t \sum_{z\in\mathcal{Z}} z q_t^{n_j}(z) \mathrm{d}t, \text{as well as}$$
$$\gamma(t) = x + \int_0^t \sum_{z\in\mathcal{Z}} z q_t(z) \mathrm{d}t,$$

for $t \in [0,T]$. Then, it follows that

$$\begin{aligned}\rho_T(\gamma^{n_j}, \gamma) &= \sup_{0\leq t\leq T} \left\| \int_0^t \sum_{z\in\mathcal{Z}} z[q_t^{n_j}(z) - q_t(z)] \mathrm{d}t \right\| \\ &\leq \sup_{0\leq t\leq T} \int_0^t \sum_{z\in\mathcal{Z}} \|z\| \left| q_t^{n_j}(z) - q_t(z) \right| \mathrm{d}t.\end{aligned}$$

Specializing eq. (37) to the case $g(t,z) = \|z\|$ shows that the upper bound converges to 0 as $j \to \infty$. This shows that the subsequence $\{\gamma^{n_j}\}$ converges to the limit function γ, which is itself absolutely continuous with almost everywhere derivative

$$\dot{\gamma}(t) = \sum_{z\in\mathcal{Z}} z q_t(z).$$

This shows that $\Gamma(M)$ is relatively compact. To show that it is compact, it remains to show that $\gamma \in \Gamma(M)$. To this end, let us define Borel measures $m^n(\mathrm{d}t \times \mathrm{d}u) = \delta_{\dot{\gamma}^n(t)}(\mathrm{d}u)\mathrm{d}t$, and $m(\mathrm{d}t \times \mathrm{d}u) = \delta_{\dot{\phi}(t)}(\mathrm{d}u)\mathrm{d}t$ on the space of Borel measures $\mathsf{M}^+([0,T] \times \mathsf{Z})$. Endowing this space with the Prohorov metric shows that $m^n \to m$ weakly in the sense of measures. Now we use the lower semi-continuity and convexity of the running costs function in the control variable to complete the proof via the following string of inequalities:

$$\begin{aligned}M &> \liminf_{n\to\infty} c^{\eta}_{x,T}(\gamma^n) = \liminf_{n\to\infty} \int_0^T L^{\eta}(\gamma^n(t), \dot{\gamma}^n(t)) \mathrm{d}t \\ &= \liminf_{n\to\infty} \int_{[0,T]\times\mathsf{Z}} L^{\eta}(\gamma^n(t), u) m^n(\mathrm{d}t \times \mathrm{d}u) \\ &\geq \int_{[0,T]\times\mathsf{Z}} L^{\eta}(\gamma(t), u) m(\mathrm{d}t \times \mathrm{d}u) \\ &= \int_0^T \left(\int_{\mathsf{Z}} L^{\eta}(\gamma(t), u) \delta_{\dot{\gamma}(t)}(\mathrm{d}u) \right) \mathrm{d}t \\ &\geq \int_0^T L^{\eta}(\gamma(t), \dot{\gamma}(t)) \mathrm{d}t.\end{aligned}$$

□

Corollary 1. *For every $x \in \mathsf{X}, \eta > 0$ and $T > 0$, the path cost functional $c^{\eta}_{x,T}(\cdot) : \mathbf{C}([0,T] : \mathsf{X}) \to [0,\infty]$ is lower semi-continuous.*

Proof. This follows from Proposition 1 and the fact that a function is lower semi-continuous if and only if it has closed level sets. □

Theorem 5.4, due to [47], shows that the sample paths of the interpolated processes $\{\hat{X}^{N,\eta}(t); t \geq 0\}$ satisfy a large deviation principle with good rate function $c^{\eta}_{x,T}$. To state this result, we use the notation $\hat{\mathbf{X}}^{N,\eta}_{[0,T]}$ as shorthand for $\{\hat{X}^{N,\eta}(t); t \in [0,T]\}$.

Theorem 5.4. ([47]) *Suppose that the processes $\{\hat{\mathbf{X}}^{N,\eta}\}_{N=N_0}^{\infty}$ have initial conditions $x^N \in \mathcal{X}^N$ satisfying $\lim_{N\to\infty} x^N = x \in X$. Let $\Gamma \subseteq \mathbf{C}([0,T] : \mathsf{X})$ be a Borel set. Then*

$$\limsup_{N\to\infty} \frac{\eta}{N} \log \mathbb{P}_{x^N}\left[\hat{\mathbf{X}}^{N,\eta}_{[0,T]} \in \Gamma\right] \leq -\inf_{\gamma\in\mathrm{cl}(\Gamma)} \eta c^{\eta}_{x,T}(\gamma), \quad \textit{and} \tag{38a}$$

$$\liminf_{N\to\infty} \frac{\eta}{N} \log \mathbb{P}_{x^N}\left[\hat{\mathbf{X}}^{N,\eta}_{[0,T]} \in \Gamma\right] \geq -\inf_{\gamma\in\mathrm{int}(\Gamma)} \eta c^{\eta}_{x,T}(\gamma). \tag{38b}$$

We refer to inequality (38a) as the *large deviation principle upper bound*, and to (38b) as the *large deviation principle lower bound*.

In [47], the LDP upper and lower bounds where used to estimate the exit time from attractors of the mean dynamics (M), and to bound the stationary distribution weights for the finite population process $\mathbf{X}^{N,\eta}$ when the noise level $\eta > 0$ is fixed. To illustrate this, let x^* denote a locally attracting equilibrium of the mean dynamics and O an open set inside the basin of attraction of x^*. Define

$$C^{\eta}(x,y) = \inf_{T>0} \inf\{c_{x,T}(\gamma) | \gamma \in \mathbf{AC}_x[0,T], \gamma(T) = y\},$$
$$C^{\eta}(\mathrm{bd}(\mathsf{O})) = \inf\{C^{\eta}(x^*,y) | y \in \mathrm{bd}(\mathsf{O})\}.$$

[47] show that for $x^N \to x \in \mathsf{O}$

$$\lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{x^N}[\tau^{N,\eta}_{\mathrm{bd}(\mathsf{O})}] = C^{\eta}(\mathrm{bd}(\mathsf{O})), \tag{39}$$

where $\hat{\tau}^{N,\eta}_D = \inf\{t \geq 0 | \hat{X}^{N,\eta}_t \in D\}$ is the first hitting time of the linearly interpolated process on the set $D \subset \mathsf{X}$. Unraveling expression (39) states that rough large population estimates of the exit time from any open set inside the basin of attraction of an equilibrium are estimated as

$$\mathbb{E}_{x^N}[\hat{\tau}^{N,\eta}_{\mathrm{bd}(\mathsf{O})}] \approx \exp\left(-NC^{\eta}(\mathrm{bd}(\mathsf{O}))\right) \text{ for } N \to \infty.$$

Estimates of this kind are folklore in the Freidlin-Wentzell theory of random perturbations of dynamical systems. Clearly, such estimates provide only rough asymptotic estimates capturing only the exponential asymptotics of characteristic statistics of the process. A more refined analysis can potentially be given via the potential-theoretic approach developed in [10]. A tractable setting where this more ambitious program could be potentially realized is the logit dynamics in potential games. [47] show that in this important special case, the variational problems involved in the large deviation estimates can be computed exactly. A key role here is played by the logit potential function

$$f^{\eta}(x) = \frac{1}{\eta} f(x) - h(x), \tag{40}$$

where $h(x) = \sum_{i=1}^{n} x_i \log(x_i)$ is the Boltzmann-Shannon entropy and $f(x)$ is a potential function for the game f, i.e. $\nabla f(x) = F(x)$ [44]. [24] show that the logit potential function is a Lyapunov function for the logit dynamics

$$\dot{x}_i = \frac{\exp(\eta^{-1}F_i(x))}{\sum_{k\in S}\exp(\eta^{-1}F_k(x))} - x_i.$$

Proposition 4 in [47] proves that if x^* is a globally attracting equilibrium of the logit dynamics, then

$$C^\eta(\text{bd}(\mathsf{O})) = \min_{y\in\text{bd}(\mathsf{O})}(f^\eta(x^*) - f^\eta(y)) \tag{41}$$

where O is an open neighborhood of x^* relative to X.

Open Problem 2. *Assume that f is a Morse function. Compute precise metastable exit times [10] for the evolutionary dynamics $\mathbf{X}^{N,\eta}$ as $N\to\infty$ and $\eta>0$ fixed.*

In principle the program described in Section 4 carries over immediately to the large population limit. Indeed, we can naturally define the radius and coradius of a point x with respect to the target set T and the source set S respectively as follows.

$$\mathsf{rad}^\eta(x,\mathsf{T}) = \inf\{c^\eta_{x,T}(\gamma)|\gamma\in\Gamma_T(\{x\},\mathsf{T}), T>0\}, \tag{42}$$

and the coradius

$$\mathsf{corad}^\eta(\mathsf{S},x) = \inf\{c^\eta_{\gamma(0),T}(\gamma)|\gamma\in\Gamma_T(\mathsf{S},\{x\}), T>0\}. \tag{43}$$

Open Problem 3. *Develop a Hamilton-Jacobi characterization as in Theorem 4.1 for $x\mapsto\mathsf{rad}^\eta(x,\mathsf{T})$ and $x\mapsto\mathsf{corad}^\eta(\mathsf{S},x)$.*

5.3. **Double limits.** By now we understand stochastic stability analysis in games in the following situations:

1. The Small Noise Limit ($\eta\to 0$);
2. The Small Noise Double Limit ($\eta\to 0$ and then $N\to\infty$);
3. The Large population limit ($N\to\infty$).

An intriguing question is the following:

Open Problem 4. *How do we quantitatively analyze stochastic stability in games in the Large Population Double Limit (LPDL) $N\to\infty$ and then $\eta\to 0$.*

There are at least two good reasons why we believe that a solution to this open problem would be a milestone. First, the small noise limit is the scenario where randomness in the players' choices becomes negligible. Intuitively, $\eta\to 0$ means that players act more and more like stage-game optimizers. Second, we have by now a quite complete understanding how to analyze stochastic evolutionary game dynamics in the small noise double limit. This limiting case is tractable, thanks to the optimal control approach pioneered in [52]. While it seems clear that the Hamilton-Jacobi approach of Theorem 4.1 carries over to the large population limit, the issue here is that we have no closed-form expression for the running cost function L^η. In fact, the only information we have about this function is that it is the Fenchel conjugate of the logarithmic moment generating function of the increment process. Except for special cases (notably logit evolution and potential games) we are not aware of any successful attempt to compute the radius and coradius when the noise is bounded away from zero. However, we know how to characterize these quantities when $\eta\to 0$. Hence, if we could prove a result saying that the (co-)radius of a stable equilibrium computed under SNDL provides a good approximation to the same quantity under the reverse order of limit, we could use the former precise results as a valid approximation for the LPDL. This is the big open question to the field which remains to be solved.

Another interesting exercise is to solve the exit and transition cost problems in the LPDL in games where the solution is known in the opposite SNDL. We propose the following two open problems in this direction.

Open Problem 5. *Can there be indirect exit and binding state constraints in the three-strategy coordination games considered by* [2] *in the large population double limit under the logit choice or probit choice rules?*

Open Problem 6. *In the large population double limit, will the presence of bilingual option alter the equilibrium chosen in the long run under the logit choice rule in the class of three-strategy bilingual games considered by* [1]*?*

The rest of this section illustrates a known instance where double limits agree. However, we also discuss that agreement should not be expected in games with more than two actions.

5.4. **Agreement of the double limits in binary choice games.** In this section, we consider evolution in two-strategy games under a general class of perturbed best response protocols (see Example 4) which includes the logit and probit choice rules. This section is based on [42].

When the population plays a game with two strategies, the state space $\mathcal{X}^N$ of the Markov process $\mathbf{X}^{N,\eta}$ is a uniformly spaced grid on the unit interval. Because agents in our model switch strategies sequentially, transitions of the process $\mathbf{X}^{N,\eta}$ are always between adjacent states, implying that it is a birth and death process. Well known results from Markov chains guarantees the existence of a unique invariant measure $\mu_x^{N,\eta}$ of $\mathbf{X}^{N,\eta}$ which summarizes the long run behavior of the Markov process. For a given two-strategy population game and a perturbed best response protocol, [42] showed that there exists a unique function $\Delta I : \mathsf{X} \to (-\infty, 0]$ such that Eqs. (44) and (45) hold.

$$\lim_{N\to\infty} \lim_{\eta\to\infty} \max_{x\in\mathcal{X}^N} \left| \frac{\eta}{N} \log \mu_x^{N,\eta} - \Delta I(x) \right| = 0 \tag{44}$$

$$\lim_{\eta\to\infty} \lim_{N\to\infty} \max_{x\in\mathcal{X}^N} \left| \frac{\eta}{N} \log \mu_x^{N,\eta} - \Delta I(x) \right| = 0 \tag{45}$$

In words, the above equations say that whether one takes the small noise limit before the large population limit, or the large population limit before the small noise limit, the rates of decay of the stationary distribution are captured by the same function. Thus when agents employ perturbed best response protocols in two-strategy population games, the choice of the order of limits has no effect on our predictions of long run behavior. In games with more than two strategies, the order of limits can matter on the long run behavior as we demonstrate this by an example in a three-strategy game.

5.5. **A 3-strategy example.** The following example seems to be the minimal one where the order of limits in stochastic stability calculus matters. The example is given by the payoff function

$$F(x_1, x_2, x_3) = \begin{pmatrix} x_1 \\ 1 - x_1 \\ 1 - x_1 \end{pmatrix}, \tag{46}$$

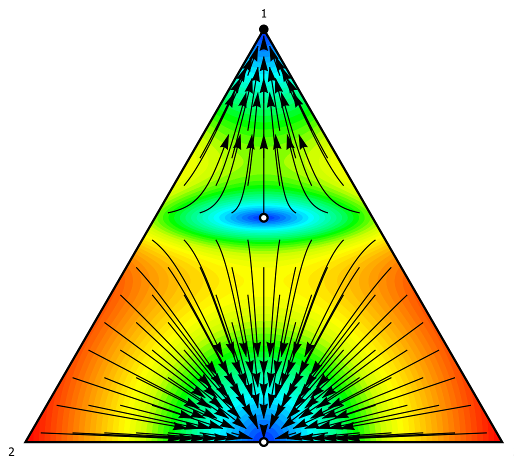


FIGURE 4. Phase portrait of the logit dynamics $V^{0.1}$ for example (46).

generated by the game matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. This game has one strict Nash equilibrium $e_1 = (1, 0, 0)^\top$, and a full component of Nash equilibria

$$\mathsf{K} = \{x \in \mathsf{X} | x_2 + x_3 = 1\}.$$

The set $\{e_1\} \cup \mathsf{K}$ is the global attractor of the best-response dynamics (see Figure 4).

Introducing small perturbations in the choice dynamics, the component is not an equilibrium anymore. Let V^η be the vector field of logit dynamics. Explicitly, this gives rise to a 2-dimensional system

$$\begin{aligned}
\dot{x}_1 &= \frac{\exp(\eta^{-1}x_1)}{\exp(\eta^{-1}x_1) + 2\exp(\eta^{-1}(1-x_1))} - x_1, \\
\dot{x}_2 &= \frac{\exp(\eta^{-1}(1-x_1))}{\exp(\eta^{-1}x_1) + 2\exp(\eta^{-1}(1-x_1))} - x_2, \\
\dot{x}_3 &= \frac{\exp(\eta^{-1}(1-x_1))}{\exp(\eta^{-1}x_1) + 2\exp(\eta^{-1}(1-x_1))} - x_3.
\end{aligned}$$

$\alpha^\eta(x) = \exp(\eta^{-1}x)$ and $\beta^\eta(x) = \exp(\eta^{-1}(1-x))$, and

$$M^\eta(x) = \frac{\alpha^\eta(x)}{\alpha^\eta(x) + 2\beta^\eta(x)} \qquad x \in [0, 1].$$

Then, the set of rest points of the mean dynamic is given by

$$\mathcal{E}^\eta = \left\{ x \in \mathsf{X} | x_1 = M^\eta(x_1), x_2 = x_3 = \frac{1 - M^\eta(x_1)}{2} \right\}.$$

For $\eta \to \infty$, the logit dynamics has a unique equilibrium at the barycenter $(1/3, 1/3, 1/3)^\top$, and a continuity argument implies that the set of solutions remains a singleton for η sufficiently large. Observe that $\mathbf{A}^\top = \mathbf{A}$, and thus F is a potential game with potential function

$$f(x) = \frac{1}{2}x^\top \mathbf{A}x = x_1^2 - x_1 + \frac{1}{2}.$$

It is then well known that the stochastic process $\mathbf{X}^{N,\eta}$ is irreducible with invariant measure

$$U^{N,\eta}(x) = \exp\left(\frac{N}{\eta}f^{N,\eta}(x)\right), \text{ where } f^{N,\eta}(x) = f(x) + \frac{\eta}{N}\ln\left(\frac{N!}{\prod_{k\in S}(Nx_k)!}\right)$$

See, e.g. [43], chapter 11, for a proof of this result, and [51] for the corresponding result for network games. The invariant distributions reads thus as

$$\mu^{N,\eta}(x) = \frac{U^{N,\eta}(x)}{\sum_{y\in\mathcal{X}^N} U^{N,\eta}(y)}.$$

For $N \to \infty$, one readily sees that $f^{N,\eta} \to \eta f^\eta$ uniformly on X, where f^η is the logit potential function defined in eq. (40). The potential function f has a global minimum when $x_1 = 1/2$. The local maxima are when $x_1 \in \{0, 1\}$. Translating to the simplex, we therefore see that

$$\begin{aligned} \mathsf{S}_1 &= \operatorname{argmin}\{f(x)|x \in \mathsf{X}\} = \{x \in \mathsf{X}|x_1 = 1/2\}, \text{ and} \\ \mathsf{S}_2 &= \operatorname{argmax}\{f(x)|x \in \mathsf{X}\} = \{e_1\} \cup \mathsf{K}. \end{aligned}$$

The separatrix of the best-response regions of the attractors is given by the hyperplane intersecting the points $\frac{1}{2}e_1 + \frac{1}{2}e_2$ and $\frac{1}{2}e_1 + \frac{1}{2}e_3$. Let us investigate qualitatively the conjectured selection patterns under the respective double limits.

- SNDL: Looking at the exact formula for the invariant measure, it is tempting to conclude that $\mu^{N,\eta}(\{e_1\}) = O(\mu^{N,\eta}(\mathsf{K}))$ for N large enough and $\eta \to 0$. This suggests that S_2 is selected in this limit case.
- LPDL: In the large population limit, for $\eta > 0$, the invariant measure is a Dirac measure located at the mean dynamics unique rest point, i.e. $\mu^\eta = \delta_{\mathcal{E}^\eta}$. The selection in the double limit is thus determined by the weak limit points of the sequence $\{\delta_{\mathcal{E}^\eta}\}_{\eta>0}$. The phase diagram suggests that the set of weak limit points is generated by the two measures δ_{e_1} and $\delta_{\frac{1}{2}e_2+\frac{1}{2}e_3}$.

These heuristic arguments suggest that the limit operations make different predicitions.

Open Problem 7. *Make these claims rigorous.*

6. **Conclusion.** A lot needs still to be done in order to understand stochastic stability calculus in games. We tried to make this clear by formulating seven open problems which we consider as important milestones left undiscovered. However, it is equally important to consider different techniques than the ones described here. Local approximations via diffusion equations might be a worthwhile complementary approach shedding some light on the transient behavior of the game dynamics. We have not pursued this direction of research at all, but some results have been derived in [41] and [60]. We believe that more research in that direction would be very

important to connect the long-run and short-run behavior of the stochastic game dynamics.

Acknowledgments. In this article we tried to give an overview about the state-of-the-art of stochastic stability analysis in games. We deliberately focused on results that where obtained in collaboration with Bill over the years. The authors, Mathias Staudigl and Srinivas Arigapudi dedicate this paper to the memory of Bill Sandholm. Thanks for teaching us this material, for long years of friendship, support and guidance. Without Bill's restless efforts, none of the results described in this overview article would have been possible. Bill's ideas will continue shaping the research agenda in this field, sadly without himself being among us. Bill, your presence is truly missed in this field and your personal and scientific contributions will never be forgotten.

REFERENCES

[1] S. Arigapudi, Exit from equilibrium in coordination games under probit choice, *Games Econom. Behav.*, **122** (2020), 168–202.
[2] S. Arigapudi, Transitions between equilibria in bilingual games under logit choice, *J. Math. Econom.*, **86** (2020), 24–34.
[3] M. Beckmann, C. B. McGuire and C. B. Winsten, *Studies in the Economics of Transportation*, Yale University Press, New Haven, CT, 1956.
[4] M. Benaïm, Recursive algorithms, urn processes, and chaining number of chain recurrent sets, *Ergodic Theory Dynam. Systems*, **18** (1998), 53–87.
[5] M. Benaïm and J. W. Weibull, Deterministic approximation of stochastic evolution in games, *Econometrica*, **71** (2003), 873–903.
[6] M. Benaïm and J. W. Weibull, *Mean-Field Approximation of Stochastic Population Processes in Games*, Université de Neuchâtel and Stockholm School of Economics, 2009. Available from: https://hal.archives-ouvertes.fr/hal-00435515/document.
[7] L. E. Blume, How noise matters, *Games Econom. Behav.*, **44** (2003), 251–271.
[8] L. E. Blume, The statistical mechanics of strategic interaction, *Games Econom. Behav.*, **5** (1993), 387–424.
[9] I. M. Bomze, Regularity versus degeneracy in dynamics, games, and optimization: A unified approach to different aspects, *SIAM Rev.*, **44** (2002), 394–414.
[10] A. Bovier and F. den Hollander, *Metastability. A Potential-Theoretic Approach*, Grundlehren der Mathematischen Wissenschaften, 351, Springer, Cham, 2015.
[11] O. Catoni, Simulated annealing algorithms and Markov chains with rare transitions, in *Séminaire de Probabilités XXXIII*, Lecture Notes in Math., 1709, Springer, Berlin, 1999, 69–119.
[12] G. C. Chasparis, Stochastic stability of perturbed learning automata in positive-utility games, *IEEE Trans. Automat. Control*, **64** (2019), 4454–4469.
[13] I.-K. Cho, N. Williams and T. J. Sargent, Escaping Nash inflation, *Rev. Econom. Stud.*, **69** (2002), 1–40.
[14] E. Dokumacı and W. H. Sandholm, Large deviations and multinomial probit choice, *J. Econom. Theory*, **146** (2011), 2151–2158.
[15] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1997.
[16] G. Ellison, Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution, *Rev. Econom. Stud.*, **67** (2000), 17–45.
[17] D. Foster and P. Young, Stochastic evolutionary game dynamics, *Theoret. Population Biol.*, **38** (1990), 219–232.
[18] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren der Mathematischen Wissenschaften, 260, Springer, New York, 1998.
[19] J. C. Harsanyi and R. Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA, 1988.

[20] S. Hart and A. Mas-Colell, Uncoupled dynamics do not lead to Nash equilibrium, *Amer. Econom. Rev.*, **93** (2003), 1830–1836.

[21] J. Hofbauer, Deterministic evolutionary game dynamics, in *Evolutionary Game Dynamics*, Proc. Sympos. Appl. Math., 69, AMS Short Course Lecture Notes, Amer. Math. Soc., Providence, RI, 2011, 61–79.

[22] J. Hofbauer, *Stability for the Best Response Dynamics*, unpublished manuscript, University of Vienna, 1995.

[23] J. Hofbauer, J. Oechssler and F. Riedel, Brown-von Neumann-Nash dynamics: The continuous strategy case, *Games Econom. Behav.*, **65** (2009), 406–429.

[24] J. Hofbauer and W. H. Sandholm, Evolution in games with randomly disturbed payoffs, *J. Econom. Theory*, **132** (2007), 47–69.

[25] J. Hofbauer and W. H. Sandholm, Survival of dominated strategies under evolutionary dynamics, *Theor. Econ.*, **6** (2011), 341–377.

[26] J. Hofbauer and K. Sigmund, Evolutionary game dynamics, *Bull. Amer. Math. Soc. (N.S.)*, **40** (2003), 479–519.

[27] J. Hofbauer and S. Sorin, Best response dynamics for continuous zero-sum games, *Discrete Contin. Dyn. Syst. Ser. B*, **6** (2006), 215–224.

[28] Y. Kifer, A discrete-time version of the Wentzell-Freidlin theory, *Ann. Probab*, **18** (1990), 1676–1692.

[29] Y. Kifer, *Random Perturbations of Dynamical Systems*, Progress in Probability and Statistics, 16, Birkhäuser Boston, Inc., Boston, MA, 1988.

[30] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*, Applications of Mathematics (New York), 35, Springer-Verlag, New York, 1997.

[31] J. R. Marden and J. S. Shamma, *Game-Theoretic Learning in Distributed Control*, Handbook of Dynamic Game Theory, Springer, Cham, 2018, 511–546.

[32] J. R. Marden, H. P. Young and L. Y. Pao, Achieving pareto optimality through distributed learning, *SIAM J. Control Optim.*, **52** (2014), 2753–2770.

[33] J. Maynard Smith and G. R. Price, The logic of animal conflict, *Nature*, **246** (1973), 15–18.

[34] P. Mertikopoulos and M. Staudigl, On the convergence of gradient-like flows with noisy gradient input, *SIAM J. Optim.*, **28** (2018), 163–197.

[35] K. Michihiro and R. Rob, Bandwagon effects and long run technology choice, *Games Econom. Behav.*, **22** (1998), 30–60.

[36] J. F. Nash Jr., *Non-Cooperative Games*, Ph.D thesis, Princeton University, 1950.

[37] D. Oyama and S. Takahashi, Contagion and uninvadability in local interaction games: The bilingual game and general supermodular games, *J. Econom. Theory*, **157** (2015), 100–127.

[38] N. Quijano, C. Ocampo-Martinez, J. Barreiro-Gomez, G. Obando, A. Pantoja and E. Mojica-Nava, The role of population games and evolutionary dynamics in distributed control systems: The advantages of evolutionary game theory, *IEEE Control Syst.*, **37** (2017), 70–97.

[39] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, 28, Princeton University Press, Princeton, NJ, 1970.

[40] R. W. Rosenthal, A class of games possessing pure-strategy Nash equilibria, *Internat. J. Game Theory*, **2** (1973), 65–67.

[41] W. H Sandholm, Evolution and equilibrium under inexact information, *Games Econom. Behav.*, **44** (2003), 343–378.

[42] W. H. Sandholm, Orders of limits for stationary distributions, stochastic dominance, and stochastic stability, *Theor. Econ.*, **5** (2010), 1–26.

[43] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, Economic Learning and Social Evolution, MIT Press, Cambridge, MA, 2010.

[44] W. H. Sandholm, Potential games with continuous player sets, *J. Econom. Theory*, **97** (2001), 81–108.

[45] W. H. Sandholm, Stochastic evolutionary game dynamics: Foundations, deterministic approximation, and equilibrium selection, in *Evolutionary Game Dynamics*, Proc. Sympos. Appl. Math., 69, AMS Short Course Lecture Notes, Amer. Math. Soc., Providence, RI, 2011, 111–141.

[46] W. H. Sandholm and M. Staudigl, Large deviations and stochastic stability in the small noise double limit, *Theor. Econ.*, **11** (2016), 279–355.

[47] W. H. Sandholm and M. Staudigl, Sample path large deviations for stochastic evolutionary game dynamics, *Math. Oper. Res.*, **43** (2018), 1348–1377.

[48] W. H. Sandholm, H. V. Tran and S. Arigapudi, Hamilton-Jacobi equations with semilinear costs and state constraints, with applications to large deviations in games, *Math. Oper. Res.*.

[49] K. Satsukawa, K. Wada and T. Iryo, Stochastic stability of dynamic user equilibrium in unidirectional networks: Weakly acyclic game approach, *Trans. Res. Part B: Methodological*, **125** (2019), 229–247.

[50] D. Shah and J. Shin, Dynamics in congestion games, in *Proceedings of the ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '10, New York, 2010, 107–118.

[51] M. Staudigl, Co-evolutionary dynamics and Bayesian interaction games, *Internat. J. Game Theory*, **42** (2013), 179–210.

[52] M. Staudigl, Potential games in volatile environments, *Games Econom. Behav.*, **72** (2011), 271–287.

[53] M. Staudigl, Stochastic stability in asymmetric binary choice coordination games, *Games Econom. Behav.*, **75** (2012), 372–401.

[54] C. Taylor, D. Fudenberg, A. Sasaki and M. A. Nowak, Evolutionary game dynamics in finite populations, *Bull. Math. Biol.*, **66** (2004), 1621–1644.

[55] P. D. Taylor and L. B. Jonker, Evolutionary stable strategies and game dynamics, *Math. Biosci.*, **40** (1978), 145–156.

[56] A. Traulsen and C. Hauert, Stochastic evolutionary game dynamics, in *Reviews of Nonlinear Dynamics and Complexity. Vol. 2*, Wiley-VCH Verlag, Weinheim, 2009, 25–61.

[57] N. Vielle, Small perturbations and stochastic games, *Israel J. Math.*, **119** (2000), 127–142.

[58] J. W. Weibull, The mass action interpretation,*J. Econom. Theory*, **69** (1996), 165–171.

[59] H. P. Young, The evolution of conventions, *Econometrica*, **61** (1993), 57–84.

[60] D. Zhou and H. Qian, Fixation, transient landscape, and diffusion dilemma in stochastic evolutionary game dynamics, *Phys. Rev. E*, **84** (2011).

Received for publication October 2020.

E-mail address: m.staudigl@maastrichtuniversity.nl
E-mail address: arigapudi@campus.technion.ac.il