Stochastic Evolutionary Game Dynamics: Foundations, Deterministic Approximation, and Equilibrium Selection

William H. Sandholm

ABSTRACT. We present a general model of stochastic evolution in games played by large populations of anonymous agents. Agents receive opportunities to revise their strategies by way of independent Poisson processes. A *revision protocol* describes how the probabilities with which an agent chooses each of his strategies depend on his current payoff opportunities and the current behavior of the population. Over finite time horizons, the population's behavior is well-approximated by a *mean dynamic*, an ordinary differential equation defined by the expected motion of the stochastic evolutionary process. Over the infinite time horizon, the population's behavior is described by the stationary distribution of the stochastic evolutionary process. If limits are taken in the population size, the level of noise in agents' revision protocols, or both, the stationary distribution may become concentrated on a small set of population states, which are then said to be *stochastically stable*. Stochastic stability analysis allows one to obtain unique predictions of very long run behavior even when the mean dynamic admits multiple locally stable states. We present a full analysis of the asymptotics of the stationary distribution in two-strategy games under noisy best protocols, and discuss extensions of this analysis to other settings.

1. Introduction

Evolutionary game theory studies the behavior of large populations of agents who repeatedly engage in anonymous strategic interactions—that is, interactions in which each agent's outcome depends not only on his own choice, but also on the distribution of others' choices. Applications range from natural selection in animal populations, to driver behavior in highway networks, to consumer choice between different technological standards, to the design of decentralized controlled systems.

In an evolutionary game model, changes in agents' behavior may be driven either by natural selection via differences in birth and death rates in biological contexts, or by the application of myopic decision rules by individual agents in

economic contexts. The resulting dynamic models can be studied using tools from the theory of dynamical systems and from the theory of stochastic processes, as well as those from stochastic approximation theory, which provides important links between the two more basic fields.

In these notes, we present a general model of stochastic evolution in large-population games, and offer a glimpse into the relevant literature by presenting a selection of basic results. In Section 2, we describe population games themselves, and offer a few simple applications. In Sections 3 and 4, we introduce our stochastic evolutionary process. To define this process, we suppose that agents receive opportunities to revise their strategies by way of independent Poisson processes. A *revision protocol* describes how the probabilities with which an agent chooses each of his strategies depend on his current payoff opportunities and the current behavior of the population. Together, a population game, a revision protocol, and a population size implicitly define the stochastic evolutionary process, a Markov process on the set of population states. In Section 4, we show that over finite time horizons, the population's behavior is well-approximated by a *mean dynamic*, an ordinary differential equation defined by the expected motion of the stochastic evolutionary process.

To describe behavior over very long time spans, we turn to an infinite-horizon analysis, in which the population's behavior is described by the stationary distribution of the stochastic evolutionary process. We begin the presentation in Section 5, which reviews the relevant definitions and results from the theory of finite-state Markov processes and presents a number of examples. In order to obtain tight predictions about very long run play, one can examine the limit of the stationary distributions as the population size grows large, the level of noise in agents' decisions becomes small, or both. The stationary distribution may then become concentrated on a small set of population states, which are said to be *stochastically stable*. Stochastic stability analysis allows one to obtain unique predictions of very long run behavior even when the mean dynamic admits multiple locally stable states. In Sections 6 and 7 we introduce the relevant definitions, and we present a full analysis of the asymptotics of the stationary distribution for the case of two-strategy games under noisy best response protocols. This analysis illustrates how the specification of the revision protocol can influence equilibrium selection results. We conclude in Section 8 by discussing extensions of our analyses of infinite-horizon behavior to more complicated strategic settings.

This presentation is based on portions of Chapters 10–12 of Sandholm (2010c), in which a complete treatment of the topics considered here can be found.

2. Population Games

We consider games played by a single population (i.e., games in which all agents play equivalent roles). We suppose that there is a unit mass of agents, each of whom chooses a pure strategy from the set $S = \{1, \ldots, n\}$. The aggregate behavior of these agents is described by a *population state*; this is an element of the simplex $X = \{x \in \mathbb{R}_+^n \colon \sum_{j \in S} x_j = 1\}$, with x_j representing the proportion of agents choosing pure strategy j. We identify a *population game* with a continuous vector-valued payoff function $F : X \to \mathbb{R}^n$. The scalar $F_i(x)$ represents the payoff to strategy i when the population state is x.

Population state x^* is a *Nash equilibrium* of F if no agent can improve his payoff by unilaterally switching strategies. More explicitly, x^* is a Nash equilibrium if

(1) $$x_i^* > 0 \text{ implies that } F_i(x) \geq F_j(x) \text{ for all } j \in S.$$

Example 2.1. In a *symmetric two-player normal form game*, each of the two players chooses a (pure) strategy from the finite set S; which we write generically as $S = \{1, \ldots, n\}$. The game's *payoffs* are described by the matrix $A \in \mathbb{R}^{n \times n}$. Entry A_{ij} is the payoff a player obtains when he chooses strategy i and his opponent chooses strategy j; this payoff does not depend on whether the player in question is called player 1 or player 2.

Suppose that the unit mass of agents are randomly matched to play the symmetric normal form game A. At population state x, the (expected) payoff to strategy i is the linear function $F_i(x) = \sum_{j \in S} A_{ij} x_j$; the payoffs to all strategies can be expressed concisely as $F(x) = Ax$. It is easy to verify that x^* is a Nash equilibrium of the population game F if and only if x^* is a symmetric Nash equilibrium of the symmetric normal form game A. ♦

While population games generated by random matching are especially simple, many games that arise in applications are not of this form.

Example 2.2. Consider the following model of highway congestion, due to Beckmann et al. (1956). A pair of towns, Home and Work, are connected by a network of *links*. To commute from Home to Work, an agent must choose a *path* $i \in S$ connecting the two towns. The payoff the agent obtains is the negation of the delay on the path he takes. The delay on the path is the sum of the delays on its constituent links, while the delay on a link is a function of the number of agents who use that link.

Population games embodying this description are known as a *congestion games*. To define a congestion game, let Φ be the collection of links in the highway network. Each strategy $i \in S$ is a route from Home to Work, and so is identified with a set of links $\Phi_i \subseteq \Phi$. Each link ϕ is assigned a *cost function* $c_\phi : \mathbb{R}_+ \to \mathbb{R}$, whose argument is link ϕ's *utilization level* u_ϕ:

$$u_\phi(x) = \sum_{i \in \rho(\phi)} x_i, \quad \text{where } \rho(\phi) = \{i \in S : \phi \in \Phi_i\}$$

The payoff of choosing route i is the negation of the total delays on the links in this route:

$$F_i(x) = - \sum_{\phi \in \Phi_i} c_\phi(u_\phi(x)).$$

Since driving on a link increases the delays experienced by other drivers on that link (i.e., since highway congestion involves *negative externalities*), cost functions in models of highway congestion are increasing; they are typically convex as well. Congestion games can also be used to model positive externalities, like the choice between different technological standards; in this case, the cost functions are decreasing in the utilization levels. ♦

3. Revision Protocols and the Stochastic Evolutionary Process

We now introduce foundations for our models of evolutionary dynamics. These foundations are built on the notion of a revision protocol, which describes both

the timing and results of agents' myopic decisions about how to continue playing
the game at hand. This approach to defining evolutionary dynamics was developed
in Björnerstedt and Weibull (1996), Weibull (1995), Hofbauer (1995), and Benaïm
and Weibull (2003), and Sandholm (2003, 2010b).

3.1. Definitions. A *revision protocol* is a map $\rho : \mathbb{R}^n \times X \to \mathbb{R}_+^{n \times n}$ that takes
the payoff vectors π and population states x as arguments, and returns nonnegative
matrices as outputs. For reasons to be made clear below, scalar $\rho_{ij}(\pi, x)$ is called
the *conditional switch rate* from strategy i to strategy j.

To move from this notion to an explicit model of evolution, let us consider a
population consisting of $N < \infty$ members. A number of the analyses to follow will
consider the limit of the present model as the population size N approaches infinity.
When the population is of size N, the set of feasible social states is the finite set
$\mathcal{X}^N = X \cap \frac{1}{N}\mathbb{Z}^n = \{x \in X : Nx \in \mathbb{Z}^n\}$, a grid embedded in the simplex X.

A revision protocol ρ, a population game F, and a population size N define a
continuous-time evolutionary process—a Markov process $\{X_t^N\}$—on the finite state
space $\mathcal{X}^N$. A one-size-fits-all description of this process is as follows. Each agent in
the society is equipped with a "stochastic alarm clock". The times between rings
of of an agent's clock are independent, each with a rate R exponential distribution.
The ringing of a clock signals the arrival of a revision opportunity for the clock's
owner. If an agent playing strategy $i \in S$ receives a revision opportunity, he switches
to strategy $j \neq i$ with probability ρ_{ij}/R. If a switch occurs, the population state
changes accordingly, from the old state x to a new state y that accounts for the
agent's change in strategy.

To describe the stochastic evolutionary process $\{X_t^N\}$ formally, it is enough to
specify its jump rates $\{\lambda_x^N\}_{x \in \mathcal{X}^N}$, which describe the exponential rates of transitions
from each state, and its transition probabilities $\{P_{xy}^N\}_{x,y \in \mathcal{X}^N}$, which describe the
probabilities that a transition starting at state x ends at state y.

If the current social state is $x \in \mathcal{X}^N$, then Nx_i of the N agents are playing
strategy $i \in S$. Since agents receive revision opportunities independently at ex-
ponential rate R, the basic properties of the exponential distribution imply that
revision opportunities arrive in the society as a whole at exponential rate NR.

When an agent playing strategy $i \in S$ receives a revision opportunity, he
switches to strategy $j \neq i$ with probability ρ_{ij}/R. Since this choice is indepen-
dent of the arrivals of revision opportunities, the probability that the next revision
opportunity goes to an agent playing strategy i who then switches to strategy j is

$$\frac{Nx_i}{N} \times \frac{\rho_{ij}}{R} = \frac{x_i \rho_{ij}}{R}.$$

This switch decreases the number of agents playing strategy i by one and increases
the number playing j by one, shifting the state by $\frac{1}{N}(e_j - e_i)$.

Summarizing this analysis yields the following observation.

OBSERVATION 3.1. *A population game F, a revision protocol ρ, a constant R,
and a population size N define a Markov process $\{X_t^N\}$ on the state space $\mathcal{X}^N$.
This process is described by some initial state $X_0^N = x_0^N$, the jump rates $\lambda_x^N = NR$,*

and the transition probabilities

$$P^N_{x,x+z} = \begin{cases} \dfrac{x_i \rho_{ij}(F(x), x)}{R} & \text{if } z = \frac{1}{N}(e_j - e_i),\ i, j \in S,\ i \neq j, \\[2ex] 1 - \displaystyle\sum_{i \in S} \sum_{j \neq i} \dfrac{x_i \rho_{ij}(F(x), x)}{R} & \text{if } z = \mathbf{0}, \\[2ex] 0 & \text{otherwise.} \end{cases}$$

3.2. Examples. In economic contexts, revision protocols of the form

$$(2) \qquad\qquad \rho_{ij}(\pi, x) = x_j r_{ij}(\pi, x)$$

are called *imitative protocols*. These protocols can be given a very simple interpretation: when an agent receives a revision opportunity, he chooses an opponent at random and observes her strategy. If our agent is playing strategy i and the opponent strategy j, the agent switches from i to j with probability proportional to r_{ij}. Notice that the value of the population share x_j is not something the agent need know; this term in (2) accounts for the agent's observing a randomly chosen opponent.

Example 3.2. Suppose that after selecting an opponent, the agent imitates the opponent only if the opponent's payoff is higher than his own, doing so in this case with probability proportional to the payoff difference:

$$\rho_{ij}(\pi, x) = x_j [\pi_j - \pi_i]_+.$$

This protocol is known as *pairwise proportional imitation*; see Helbing (1992) and Schlag (1998). ◆

Additional references on imitative protocols include Björnerstedt and Weibull (1996), Weibull (1995), and Hofbauer (1995).

Protocols of form (2) also appear in biological contexts, starting with the work of Moran (1962), and revisited more recently by Nowak et al. (2004) among others, see Nowak (2006) and Traulsen and Hauert (2009) for further references. In these cases we refer to (2) as a *natural selection protocol*. The biological interpretation of (2) supposes that each agent is programmed to play a single pure strategy. An agent who receives a revision opportunity dies, and is replaced through asexual reproduction. The reproducing agent is a strategy j player with probability $\rho_{ij}(\pi, x) = x_j \hat{\rho}_{ij}(\pi, x)$, which is proportional both to the number of strategy j players and to some function of the prevalences and fitnesses of all strategies. Note that this interpretation requires the restriction

$$\sum_{j \in S} \rho_{ij}(\pi, x) \equiv 1.$$

Example 3.3. Suppose that payoffs are always positive, and let

$$(3) \qquad\qquad \rho_{ij}(\pi, x) = \frac{x_j\, \pi_j}{\sum_{k \in S} x_k\, \pi_k}.$$

Understood as a natural selection protocol, (3) says that the probability that the reproducing agent is a strategy j player is proportional to $x_j \pi_j$, the aggregate fitness of strategy j players.

In economic contexts, we can interpret (3) as an imitative protocol based on repeated sampling. When an agent's clock rings he chooses an opponent at random. If the opponent is playing strategy j, the agent imitates him with probability

proportional to π_j. If the agent does not imitate this opponent, he draws a new opponent at random and repeats the procedure. ♦

In the previous examples, only strategies currently in use have any chance of being chosen by a revising agent (or of being the programmed strategy of the newborn agent). Under other protocols, agents' choices are not mediated through the population's current behavior, except indirectly via the effect of behavior on payoffs. These *direct protocols* require agents to directly evaluate the payoffs of each strategy, rather than to indirectly evaluate them as under an imitative procedure.

Example 3.4. Suppose that choices are made according to the *logit choice rule*:

$$(4) \qquad \rho_{ij}(\pi, x) = \frac{\exp(\eta^{-1}\pi_j)}{\sum_{k \in S} \exp(\eta^{-1}\pi_k)}.$$

The interpretation of this protocol is simple. Revision opportunities arrive at unit rate. When an opportunity is received by an i player, he switches to strategy j with probability $\rho_{ij}(\pi, x)$, which is proportional to an exponential function of strategy j's payoffs. The parameter $\eta > 0$ is called the *noise level*. If η is large, choice probabilities under the logit rule are nearly uniform. But if η is near zero, choices are optimal with probability close to one, at least when the difference between the best and second best payoff is not too small. ♦

4. Finite Horizon Deterministic Approximation

4.1. Mean Dynamics. A revision protocol ρ, a population game F, and a population size N define a Markov process $\{X_t^N\}$ on the finite state space $\mathcal{X}^N$. We now derive a deterministic process—the *mean dynamic*—that describes the expected motion of $\{X_t^N\}$. In Section 4.3, we will describe formally the sense in which this deterministic process provides a very good approximation of the behavior of the stochastic process $\{X_t^N\}$, at least over finite time horizons and for large population sizes. But having noted this result, we will focus in this section on the deterministic process itself.

To compute the expected increment of $\{X_t^N\}$ over the next dt time units, recall first that each of the N agents receives revision opportunities via a rate R exponential distribution, and so expects to receive $R\,dt$ opportunities during the next dt time units. If the current state is x, the expected number of revision opportunities received by agents currently playing strategy i is approximately $Nx_i\,R\,dt$. Since an i player who receives a revision opportunity switches to strategy j with probability ρ_{ij}/R, the expected number of such switches during the next dt time units is approximately $Nx_i\,\rho_{ij}\,dt$. Therefore, the expected change in the number of agents choosing strategy i during the next dt time units is approximately

$$(5) \qquad N \left(\sum_{j \in S} x_j \rho_{ji}(F(x), x) - x_i \sum_{j \in S} \rho_{ij}(F(x), x) \right) dt.$$

Dividing expression (5) by N and eliminating the time differential dt yields a differential equation for the rate of change in the *proportion* of agents choosing strategy i:

$$(M) \qquad \dot{x}_i = \sum_{j \in S} x_j \rho_{ji}(F(x), x) - x_i \sum_{j \in S} \rho_{ij}(F(x), x).$$

Equation (M) is the *mean dynamic* (or *mean field*) generated by revision protocol ρ in population game F. The first term in (M) captures the inflow of agents to strategy i from other strategies, while the second captures the outflow of agents to other strategies from strategy i.

4.2. Examples. We now revisit the revision protocols from Section 3.2. To do so, we let

$$\overline{F}(x) = \sum_{i \in S} x_i F_i(x)$$

denote the *average payoff* obtained by the members of the population, and define the *excess payoff* to strategy i,

$$\hat{F}_i(x) = F_i(x) - \overline{F}(x),$$

to be the difference between strategy i's payoff and the population's average payoff.

Example 4.1. In Example 3.2, we introduced the pairwise proportional imitation protocol $\rho_{ij}(\pi, x) = x_j[\pi_j - \pi_i]_+$. This protocol generates the mean dynamic

$$(6) \qquad \dot{x}_i = x_i \hat{F}_i(x).$$

Equation (6) is the *replicator dynamic* of Taylor and Jonker (1978), the best-known dynamic in evolutionary game theory. Under this dynamic, the percentage growth rate $\dot{x}_i/x_i$ of each strategy currently in use is equal to that strategy's current excess payoff; unused strategies always remain so. There are a variety of revision protocols other than pairwise proportional imitation that generate the replicator dynamic as their mean dynamics; see Björnerstedt and Weibull (1996) and Hofbauer (1995). ♦

Example 4.2. In Example 3.3, we assumed that payoffs are always positive, and introduced the protocol $\rho_{ij}(\pi, x) \propto x_j \pi_j$, which we interpreted both as a model of biological natural selection and as a model of imitation with repeated sampling. The resulting mean dynamic,

$$(7) \qquad \dot{x}_i = \frac{x_i F_i(x)}{\sum_{k \in S} x_k F_k(x)} - x_i = \frac{x_i \hat{F}_i(x)}{\overline{F}(x)},$$

is the *Maynard Smith replicator dynamic*, due to Maynard Smith (1982). This dynamic only differs from the standard replicator dynamic (6) by a change of speed, with motion under (7) being relatively fast when average payoffs are relatively low. In multipopulation models, the two dynamics are less similar because the changes in speed may differ across populations, affecting the direction of motion. ♦

Example 4.3. In Example 3.4 we introduced the logit choice rule $\rho_{ij}(\pi, x) \propto \exp(\eta^{-1}\pi_j)$. The corresponding mean dynamic,

$$(8) \qquad \dot{x}_i = \frac{\exp(\eta^{-1}F_i(x))}{\sum_{k \in S} \exp(\eta^{-1}F_k(x))} - x_i,$$

is called the *logit dynamic*, due to Fudenberg and Levine (1998). ♦

We summarize these and other examples of revision protocols and mean dynamics in Table 1. Dynamics from the table that have not been mentioned so far include the best response dynamic of Gilboa and Matsui (1991), the BNN dynamic of Brown and von Neumann (1950), and the Smith (1984) dynamic. Discussion, examples,

Revision protocol	Mean dynamic	Name
$\rho_{ij} = x_j[\pi_j - \pi_i]_+$	$\dot{x}_i = x_i \hat{F}_i(x)$	replicator
$\rho_{ij} = \dfrac{\exp(\eta^{-1}\pi_j)}{\sum_{k \in S} \exp(\eta^{-1}\pi_k)}$	$\dot{x}_i = \dfrac{\exp(\eta^{-1}F_i(x))}{\sum_{k \in S} \exp(\eta^{-1}F_k(x))} - x_i$	logit
$\rho_{ij} = 1_{\{j = \operatorname{argmax}_{k \in S} \pi_k\}}$	$\dot{x} \in B^F(x) - x$	best response
$\rho_{ij} = [\pi_j - \sum_{k \in S} x_k \pi_k]_+$	$\dot{x}_i = [\hat{F}_i(x)]_+ - x_i \sum_{j \in S} [\hat{F}_j(x)]_+$	BNN
$\rho_{ij} = [\pi_j - \pi_i]_+$	$\dot{x}_i = \sum_{j \in S} x_j[F_i(x) - F_j(x)]_+$ $-x_i \sum_{j \in S} [F_j(x) - F_i(x)]_+$	Smith

Table 1: Five basic deterministic dynamics.

and results concerning these and other deterministic dynamics can be found in J. Hofbauer's contribution to this volume.

4.3. Deterministic Approximation Theorem. In Section 3, we defined the Markovian evolutionary process $\{X_t^N\}$ from a revision protocol ρ, a population game F, and a finite population size N. In Section 4.1, we argued that the expected motion of this process is captured by the mean dynamic

$$(M) \qquad \dot{x}_i = V^F(x) = \sum_{j \in S} x_j \rho_{ji}(F(x), x) - x_i \sum_{j \in S} \rho_{ij}(F(x), x).$$

The basic link between the Markov process $\{X_t^N\}$ and its mean dynamic (M) is provided by the following theorem (Kurtz (1970), Sandholm (2003), Benaïm and Weibull (2003)).

THEOREM 4.4 (Deterministic Approximation of $\{X_t^N\}$). *Suppose that V^F is Lipschitz continuous. Let the initial conditions $X_0^N = x_0^N$ converge to state $x_0 \in X$, and let $\{x_t\}_{t \geq 0}$ be the solution to the mean dynamic* (M) *starting from x_0. Then for all $T < \infty$ and $\varepsilon > 0$,*

$$\lim_{N \to \infty} \mathbb{P}\left(\sup_{t \in [0,T]} \left|X_t^N - x_t\right| < \varepsilon\right) = 1.$$

Thus, when the population size N is large, nearly all sample paths of the Markov process $\{X_t^N\}$ stay within ε of a solution of the mean dynamic (M) through time T. By choosing N large enough, we can ensure that with probability close to one, X_t^N and x_t differ by no more than ε for all t between 0 and T (Figure 1).

The intuition for this result comes from the law of large numbers. At each revision opportunity, the increment in the process $\{X_t^N\}$ is stochastic. Still, the expected number of revision opportunities that arrive during the brief time interval $I = [t, t + dt]$ is large—in particular, of order $N\,dt$. Since each opportunity leads to an increment of the state of size $\frac{1}{N}$, the size of the overall change in the state during time interval I is of order dt. Thus, during this interval there are a large number of revision opportunities, each following nearly the same transition probabilities,
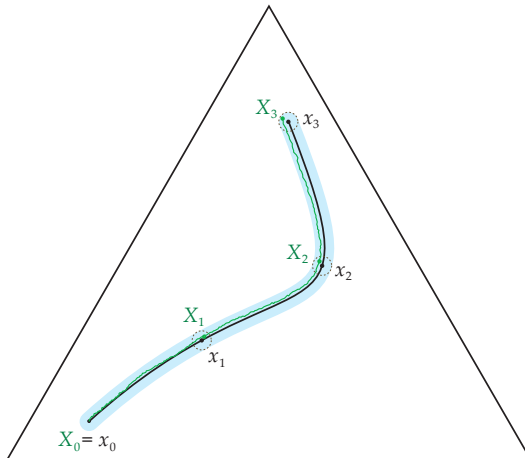
Figure 1: Deterministic approximation of the Markov process $\{X_t^N\}$.

and hence having nearly the same expected increments. The law of large numbers therefore suggests that the change in $\{X_t^N\}$ during this interval should be almost completely determined by the expected motion of $\{X_t^N\}$, as described by the mean dynamic (M).

It should be emphasized that Theorem 4.4 cannot be extended to an infinite horizon result. To see why not, consider the logit choice protocol (Example 4.3), under which switches between all pairs of strategies occur with positive probability regardless of the current state. As we discuss in Section 5, this property implies that the induced Markov process $\{X_t^N\}$ is irreducible, and hence that every state in $\mathcal{X}^N$ is visited infinitely often with probability one. This fact clearly precludes an infinite horizon analogue of Theorem 4.4. However, the failure of this result introduces a new possibility, that of obtaining unique predictions of infinite horizon behavior. We consider this question in Sections 6 and 7.

4.4. Analysis of Deterministic Dynamics. With this justification in hand, one can use methods from dynamical systems theory to study the behavior of the mean dynamic (M). A large literature has considered this question for a wide range of choices of the revision protocol ρ and the game F, proving a variety of results about local stability of equilibrium, global convergence to equilibrium, and nonconvergence. Other contributions to this volume, in particular those of J. Hofbauer and R. Cressman, address such results; for general references, see Hofbauer and Sigmund (1988, 1998, 2003), Weibull (1995), Sandholm (2009), and chapters 4–9 of Sandholm (2010c).

5. Stationary Distributions

Theorem 4.4 shows that over finite time spans, the stochastic evolutionary process $\{X_t^N\}$ follows a nearly deterministic path, closely shadowing a solution trajectory of the corresponding mean dynamic (M). But if we look at longer time spans—that is, if we fix the population size N of interest and consider the position of the process at large values of t—the random nature of the process must assert itself.

If the process is generated by a *full support revision protocol*, one that always assigns positive probabilities to transitions to all neighboring states in $\mathcal{X}^N$, then $\{X_t^N\}$ must visit all states in $\mathcal{X}^N$ infinitely often. Evidently, an *infinite* horizon approximation theorem along the lines of Theorem 4.4 cannot hold. To make predictions about play over very long time spans, we need new techniques for characterizing the infinite horizon behavior of the stochastic evolutionary process. We do so by considering the *stationary distribution* μ^N of the process $\{X_t^N\}$. A stationary distribution is defined by the property that a process whose initial condition is described by this distribution will continue to be described by this distribution at all future times. If $\{X_t^N\}$ is generated by a full support revision protocol, then its stationary distribution μ^N is not only unique, but also describes the infinite horizon behavior of $\{X_t^N\}$ regardless of this process's initial distribution. In principle, this fact allows us to use the stationary distribution to form predictions about a population's very long run behavior that do not depend on its initial behavior. This contrasts sharply with predictions based on the mean dynamic (M), which generally require knowledge of the initial state.

5.1. Full Support Revision Protocols. To introduce the possibility of unique infinite-horizon predictions, we now assume in addition that the conditional switch rates are bounded away from zero: there is a positive constant $\underline{R}$ such that

$$(9) \qquad \rho_{ij}(F(x), x) \geq \underline{R} \text{ for all } i, j \in S \text{ and } x \in X.$$

We refer to a revision protocol that satisfies condition (9) as having *full support*.

Example 5.1. Best response with mutations. Under *best response with mutations* at *mutation rate* $\varepsilon > 0$, called $BRM(\varepsilon)$ for short, a revising agent switches to his current best response with probability $1 - \varepsilon$, but chooses a strategy uniformly at random (or *mutates*) with probability $\varepsilon > 0$. Thus, if the game has two strategies, each yielding different payoffs, a revising agent will choose the optimal strategy with probability $1 - \frac{\varepsilon}{2}$ and will choose the suboptimal strategy with probability $\frac{\varepsilon}{2}$. (Kandori et al. (1993), Young (1993)) ♦

Example 5.2. Logit choice. In Example 3.4 we introduced the logit choice protocol with noise level $\eta > 0$. Here we rewrite this protocol as

$$(10) \qquad \rho_{ij}(\pi) = \frac{\exp(\eta^{-1}(\pi_j - \pi_{k^*}))}{\sum_{k \in S} \exp(\eta^{-1}(\pi_k - \pi_{k^*}))}.$$

where k^* is an optimal strategy under π. Then as η approaches zero, the denominator of (10) converges to a constant (namely, the number of optimal strategies under π), so as η^{-1} approaches infinity, $\rho_{ij}(\pi, x)$ vanishes at exponential rate $\pi_{k^*} - \pi_j$. (Blume (1993, 1997)) ♦

As their noise parameters approach zero, both the BRM and logit protocols come to resemble an exact best response protocol. But this similarity masks a fundamental qualitative difference between the two protocols. Under best response with mutations, the probability of choosing a particular suboptimal strategy is independent of the payoff consequences of doing so: mutations do not favor alternative strategies with higher payoffs over those with lower payoffs. In contrast, since the logit protocol is defined using payoff perturbations that are symmetric across strategies, more costly "mistakes" are less likely to be made. One might expect the precise specification of mistake probabilities to be of little consequence.

But as we shall see below, predictions of infinite horizon behavior hinge on the relative probabilities of rare events, so that seemingly minor differences in choice probabilities can lead to entirely different predictions of behavior.

5.2. Review: Irreducible Markov Processes. The full support assumption (9) ensures that at each revision opportunity, every strategy in S has a positive probability of being chosen by the revising agent. Therefore, there is a positive probability that the process $\{X_t^N\}$ will transit from any given current state x to any other given state y within a finite number of periods. A Markov process with this property is said to be *irreducible*. Below we review some basic results about infinite horizon behavior of irreducible Markov processes on a finite state space; for details, see e.g. Norris (1997).

Suppose that $\{X_t\}_{t \geq 0}$ is an irreducible Markov process on the finite state space $\mathcal{X}$, where the process has equal jump rates $\lambda_x \equiv l$ and transition matrix P. Then there is a unique probability vector $\mu \in \mathbb{R}_+^{\mathcal{X}}$ satisfying

$$(11) \qquad \sum_{x \in \mathcal{X}} \mu_x P_{xy} = \mu_y \text{ for all } y \in \mathcal{X}.$$

The vector μ is called the *stationary distribution* of the process $\{X_t\}$. Equation (11) tells us that if we run the process $\{X_t\}$ from initial distribution μ, then at the random time of the first jump, the distribution of the process is also μ. Moreover, if we use the notation $\mathbb{P}_\pi(\cdot)$ to represent $\{X_t\}$ being run from initial distribution π, then

$$(12) \qquad \mathbb{P}_\mu (X_t = x) = \mu_x \text{ for all } x \in \mathcal{X} \text{ and } t \geq 0.$$

In other words, if the process starts off in its stationary distribution, it remains in this distribution at all subsequent times t.

While equation (12) tells us what happens if $\{X_t\}$ starts off in its stationary distribution, our main interest is in what happens to this process in the very long run if it starts in an arbitrary initial distribution π. Then as t grows large, the time t distribution of $\{X_t\}$ converges to μ:

$$(13) \qquad \lim_{t \to \infty} \mathbb{P}_\pi(X_t = x) = \mu_x \text{ for all } x \in \mathcal{X}.$$

Thus, looking at the process $\{X_t\}$ from the ex ante point of view, the probable locations of the process at sufficiently distant future times are essentially determined by μ.

To describe long run behavior from an ex post point of view, we need to consider the behavior of the process's sample paths. Here again, the stationary distribution plays the central role. Then along almost every sample path, the proportion of time spent at each state in the long run is described by μ:

$$(14) \qquad \mathbb{P}_\pi \left(\lim_{T \to \infty} \frac{1}{T} \int_0^T 1_{\{X_t = x\}} \, dt = \mu_x \right) = 1 \text{ for all } x \in \mathcal{X}.$$

We can also summarize equation (14) by saying that the limiting empirical distribution of $\{X_t\}$ is almost surely equal to μ.

In general, computing the stationary distribution of a Markov process means finding an eigenvector of a matrix, a task that is computationally daunting unless the state space, and hence the dimension of the matrix, is small. But there is a special class of Markov processes whose stationary distributions are easy to compute.

A constant jump rate Markov process $\{X_t\}$ is said to be *reversible* if it admits a *reversible distribution*: a probability distribution μ on $\mathcal{X}$ that satisfies the *detailed balance conditions*:

$$(15) \qquad\qquad \mu_x P_{xy} = \mu_y P_{yx} \text{ for all } x, y \in \mathcal{X}.$$

A process satisfying this condition is called reversible because, probabilistically speaking, it "looks the same" whether time is run forward or backward. Since summing the equality in (15) over x yields condition (11), a reversible distribution is also a stationary distribution.

  While in general reversible Markov processes are rather special, we now introduce one important case in which reversibility is ensured. A constant jump rate Markov process $\{X_t^N\}$ on the state space $\mathcal{X}^N = \{0, \frac{1}{N}, \ldots, 1\}$ is a *birth and death process* if the only positive probability transitions move one step to the right, move one step to the left, or remain still. This implies that there are vectors $p^N, q^N \in \mathbb{R}^{\mathcal{X}^N}$ with $p_1^N = q_0^N = 0$ such that the transition matrix of $\{X_t^N\}$ takes the form

$$P_{\chi y}^N \equiv \begin{cases} p_\chi^N & \text{if } y = \chi + \frac{1}{N}, \\ q_\chi^N & \text{if } y = \chi - \frac{1}{N}, \\ 1 - p_\chi^N - q_\chi^N & \text{if } y = \chi, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, the process $\{X_t^N\}$ is irreducible if $p_\chi^N > 0$ for $\chi < 1$ and $q_\chi^N > 0$ for $\chi > 0$, as we henceforth assume. For the transition matrix above, the reversibility conditions (15) reduce to

$$\mu_\chi^N q_\chi^N = \mu_{\chi-1/N}^N p_{\chi-1/N}^N \quad \text{for } \chi \in \{\tfrac{1}{N}, \ldots, 1\}.$$

Applying this formula inductively, we find that the stationary distribution of $\{X_t^N\}$ satisfies

$$(16) \qquad\qquad \frac{\mu_\chi^N}{\mu_0^N} = \prod_{j=1}^{N\chi} \frac{p_{(j-1)/N}^N}{q_{j/N}^N} \quad \text{for } \chi \in \{\tfrac{1}{N}, \ldots, 1\},$$

with μ_0 determined by the requirement that the weights in μ^N must sum to 1.

5.3. Stationary Distributions for Two-Strategy Games.
When the population plays a game with just two strategies, the state space $\mathcal{X}^N$ is a grid in the simplex in $\mathbb{R}^2$. In this case it is convenient to identify state x with the weight $\chi \equiv x_1$ that it places on strategy 1. Under this notational device, the state space of the Markov process $\{X_t^N\}$ becomes $\mathcal{X}^N = \{0, \frac{1}{N}, \ldots, 1\}$, a uniformly-spaced grid in the unit interval. We will also write $F(\chi)$ for $F(x)$ and $\rho(\pi, \chi)$ for $\rho(\pi, x)$ whenever it is convenient to do so.

  Because agents in our model switch strategies sequentially, transitions of the process $\{X_t^N\}$ are always between adjacent states, implying that $\{X_t^N\}$ is a birth and death processes. Let us now use formula (16) to compute the stationary distribution of our stochastic evolutionary process, maintaining the assumption that the process is generated by a full support revision protocol. Referring back to Section 3.1, we find that the process $\{X_t^N\}$ has constant jump rates $\lambda_\chi^N = NR$, and that

its upward and downward transition probabilities are given by

$$(17) \qquad\qquad p_{\chi}^N = (1 - \chi) \cdot \tfrac{1}{R} \rho_{01}(F(\chi), \chi) \text{ and}$$

$$(18) \qquad\qquad q_{\chi}^N = \chi \cdot \tfrac{1}{R} \rho_{10}(F(\chi), \chi).$$

Substituting formulas (17) and (18) into equation (16), we see that for $\chi \in \{\tfrac{1}{N}, \tfrac{2}{N}, \ldots, 1\}$, we have

$$\frac{\mu_{\chi}^N}{\mu_0^N} = \prod_{j=1}^{N\chi} \frac{p_{(j-1)/N}^N}{q_{j/N}^N} = \prod_{j=1}^{N\chi} \frac{(1 - \frac{j-1}{N})}{\frac{j}{N}} \cdot \frac{\frac{1}{R}\rho_{01}(F(\frac{j-1}{N}), \frac{j-1}{N})}{\frac{1}{R}\rho_{10}(F(\frac{j}{N}), \frac{j}{N})}.$$

Simplifying this expression yields the following result.

THEOREM 5.3. *Suppose that a population of N agents plays the two-strategy game F using the full support revision protocol ρ. Then the stationary distribution for the evolutionary process $\{X_t^N\}$ on $\mathcal{X}^N$ is given by*

$$\frac{\mu_{\chi}^N}{\mu_0^N} = \prod_{j=1}^{N\chi} \frac{(N - j + 1)}{j} \cdot \frac{\rho_{01}(F(\frac{j-1}{N}), \frac{j-1}{N})}{\rho_{10}(F(\frac{j}{N}), \frac{j}{N})} \quad \text{for } \chi \in \{\tfrac{1}{N}, \tfrac{2}{N}, \ldots, 1\},$$

with μ_0^N determined by the requirement that $\sum_{\chi \in \mathcal{X}^N} \mu_{\chi}^N = 1$.

In what follows, we will use Theorem 5.3 to understand the infinite-horizon behavior of the process $\{X_t^N\}$, in particular as various parameters are taken to their limiting values.

5.4. Examples. The power of infinite horizon analysis lies in its ability to generate unique predictions of play even in games with multiple strict equilibria. We now illustrate this idea by computing some stationary distributions for two-strategy coordination games under the BRM and logit rules. In all cases, we find that these distributions place most of their weight near a single equilibrium. But we also find that the two rules need not select the same equilibrium.

Example 5.4. Stag Hunt. The symmetric normal form coordination game

$$A = \begin{pmatrix} h & h \\ 0 & s \end{pmatrix}$$

with $s > h > 0$ is known as *Stag Hunt*. By way of interpretation, we imagine that each agent in a match must decide whether to hunt for hare or for stag. Hunting for hare ensures a payoff of h regardless of the match partner's choice. Hunting for stag can generate a payoff of $s > h$ if the opponent does the same, but results in a zero payoff otherwise. Each of the two strategies has distinct merits. Coordinating on Stag yields higher payoffs than coordinating on Hare. But the payoff to Hare is certain, while the payoff to Stag depends on the choice of one's partner.

Suppose that a population of agents is repeatedly matched to play Stag Hunt. If we let χ denote the proportion of agents playing Stag, then with our usual abuse of notation, the payoffs in the resulting population game are $F_H(\chi) = h$ and $F_S(\chi) = s\chi$. This population game has three Nash equilibria: the two pure equilibria, and the mixed equilibrium $\chi^* = \tfrac{h}{s}$. We henceforth suppose that $h = 2$ and $s = 3$, so that the mixed equilibrium places mass $\chi^* = \tfrac{2}{3}$ on Stag.
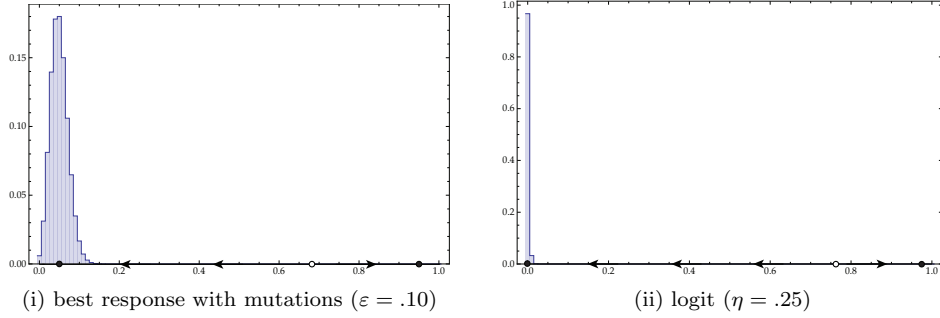
(i) best response with mutations ($\varepsilon = .10$)   (ii) logit ($\eta = .25$)

Figure 2: Stationary distribution weights μ_χ for Stag Hunt ($h = 2$, $s = 3$, $N = 100$).

Suppose that agents follow the best response with mutations protocol, with mutation rate $\varepsilon = .10$. The resulting mean dynamic,

$$\dot{\chi} = \begin{cases} \frac{\varepsilon}{2} - \chi & \text{if } \chi < \frac{2}{3}, \\ (1 - \frac{\varepsilon}{2}) - \chi & \text{if } \chi > \frac{2}{3}, \end{cases}$$

has stable rest points at $\chi = .05$ and $\chi = .95$. The basins of attraction of these rest points meet at the mixed equilibrium $\chi^* = \frac{2}{3}$. Note that the rest point that approximates the all-Hare equilibrium has the larger basin of attraction.

In Figure 2(i), we present this mean dynamic underneath the stationary distribution μ^N for $N = 100$, which we computed using the formula derived in Theorem 5.3. While the mean dynamic has two stable equilibria, nearly all of the mass in the stationary distribution is concentrated at states where between 88 and 100 agents choose Hare. Thus, while coordinating on Stag is efficient, the "safe" strategy Hare is selected by the stochastic evolutionary process.

Suppose instead that agents use the logit rule with noise level $\eta = .25$. The mean dynamic is then the logit dynamic,

$$\dot{\chi} = \frac{\exp(3\chi\eta^{-1})}{\exp(2\eta^{-1}) + \exp(3\chi\eta^{-1})} - \chi,$$

which has stable rest points at $\chi = .0003$ and $\chi = .9762$, and an unstable rest point at $\chi = .7650$, so that the basin of attraction of the "almost all-Hare" rest point $\chi = .0003$ is even larger than under BRM. Examining the resulting stationary distribution (Figure 2(ii)), we see that virtually all of its mass is placed on states where either 99 or 100 agents choose Hare, in rough agreement with the result for the BRM(.10) rule. ♦

Why does most of the mass in the stationary distribution becomes concentrated around a single equilibrium? The stochastic evolutionary process $\{X_t^N\}$ typically moves in the direction indicated by the mean dynamic. If the process begins in the basin of attraction of a rest point or other attractor of this dynamic, then the initial period of evolution generally results in convergence to and lingering near this locally stable set.

However, since BRM and logit choice lead to irreducible evolutionary processes, this cannot be the end of the story. Indeed, we know that the process $\{X_t^N\}$
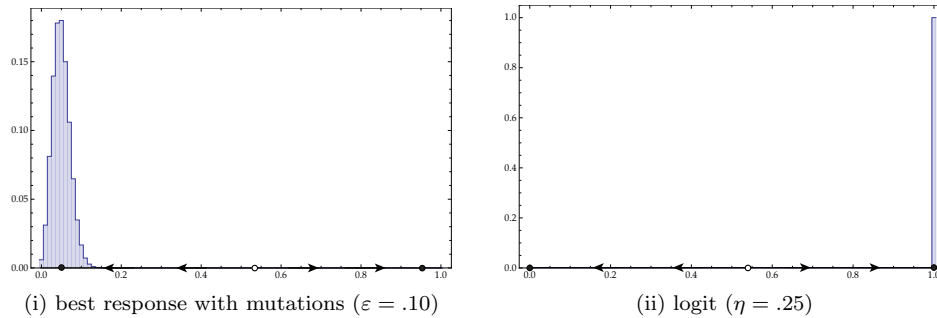
(i) best response with mutations ($\varepsilon = .10$)   (ii) logit ($\eta = .25$)

Figure 3: Stationary distribution weights μ_χ for a nonlinear Stag Hunt ($h = 2$, $s = 7$, $N = 100$).

eventually reaches all states in $\mathcal{X}^N$; in fact, it visits all states in $\mathcal{X}^N$ infinitely often. This means that the process at some point must leave the basin of the stable set visited first; it then enters the basin of a new stable set, at which point it is extremely likely to head directly the set itself. The evolution of the process continues in this fashion, with long periods of visits to each attractor punctuated by sudden jumps between the stable set.

Which states are visited most often over the infinite horizon is determined by the *relative* unlikelihoods of these rare but inevitable transitions between stable sets. In the examples above, the transitions from the Stag rest point to the Hare rest point and from the Hare rest point to the Stag rest point are both very unlikely events. But for purposes of determining the stationary distribution, what matters is that in relative terms, the former transitions are much more likely than the latter. This enables us to conclude that over very long time spans, the evolutionary process will spend most periods at states where most agents play Hare.

Example 5.5. A nonlinear Stag Hunt. We now consider a version of the Stag Hunt game in which payoffs depend nonlinearly on the population state. With our usual abuse of notation, we define payoffs in this game by $F_H(\chi) = h$ and $F_S(\chi) = s\chi^2$, with χ representing the proportion of agents playing Stag. The population game F has three Nash equilibria: the pure equilibria $\chi = 0$ and $\chi = 1$, and the mixed equilibrium $\chi^* = \sqrt{h/s}$. We focus on the case in which $h = 2$ and $s = 7$, so that $\chi^* = \sqrt{2/7} \approx .5345$.

Suppose first that a population of 100 agents play this game using the BRM(.10) rule. In Figure 3(i) we present the resulting mean dynamic beneath a graph of the stationary distribution μ^{100}. The mean dynamic has rest points at $\chi = .05$, $\chi = .95$, and $\chi^* \approx .5345$, so the the "almost all Hare" rest point again has the larger basin of attraction. As was true in the linear Stag Hunt from Example 5.4, the stationary distribution generated by the BRM(.10) rule in this nonlinear Stag Hunt places nearly all of its mass on states where at least 88 agents choose Hare.

Figure 3(ii) presents the mean dynamic and the stationary distribution μ^{100} for the logit rule with $\eta = .25$. The rest points of the logit(.25) dynamic are $\chi = .0003$, $\chi = 1$, and $\chi = .5398$, so the "almost all Hare" rest point once again has the larger

basin of attraction. Nevertheless, the stationary distribution μ^{100} places virtually all of its mass on the state in which all 100 agents choose Stag.

To summarize, our prediction for very long run behavior under the BRM(.10) rule is inefficient coordination on Hare, while our prediction under the logit(.25) rule is efficient coordination on Stag. ♦

For the intuition behind this discrepancy in predictions, recall the discussion from Section 5.1 about the basic distinction between the logit and BRM protocols: under logit choice, the probability of a "mistake" depends on its payoff consequences, while under BRM, it does not. The latter observation implies that under BRM, the probabilities of escaping from the basins of attraction of stable sets, and hence the identities of the states predominate in the very long run, depend only on the size and the shapes of the basins. In the current one-dimensional example, these shapes are always line segments, so that only the size of the basins matters; since the "almost all-Hare" state has the larger basin, it is selected under the BRM rule.

On the contrary, the probability of escaping a stable equilibrium under logit choice depends not only on the shape and size of its basin, but also on the payoff differences that must be overcome during the journey. In the nonlinear Stag Hunt game, the basin of the "almost all-Stag" equilibrium is smaller than that of the all-Hare equilibrium. But because the payoff advantage of Stag over Hare in the former's basin tends to be much larger than the payoff advantage of Hare over Stag in the latter's, it is more difficult for the population to escape the all-Stag equilibrium than the all-Hare equilibrium; as a result, the population spends virtually all periods coordinating on Stag over the infinite horizon.

We can compare the process of escaping from the basin of a stable rest point to an attempt to swim upstream. Under BRM, the strength of the stream's flow is constant, so the difficulty of a given excursion is proportional to distance. Under logit choice, the strength of the stream's flow is variable, so the difficulty of an excursion depends on how this strength varies over the distance travelled. In general, the probability of escaping from a stable set is determined by both the distance that must be travelled and the strength of the oncoming flow.

To obtain unique predictions of infinite horizon behavior, it is generally enough either that the population size not be too small, or that the noise level in agents' choices not be too large. But one can obtain cleaner and more general results by studying the limiting behavior of the stationary distribution as the population size approaches infinity, the noise level approaches zero, or both. This approach to studying infinite horizon behavior, known as *stochastic stability theory*.

One difficulty that can arise in this setting is that the prediction of infinite horizon behavior can depend on the identity or on the order in which limits are taken. Our last example, based on Binmore and Samuelson (1997), illustrates this point.

Example 5.6. Consider a population of agents who are matched to play the symmetric normal form game with strategy set $S = \{0, 1\}$ and payoff matrix

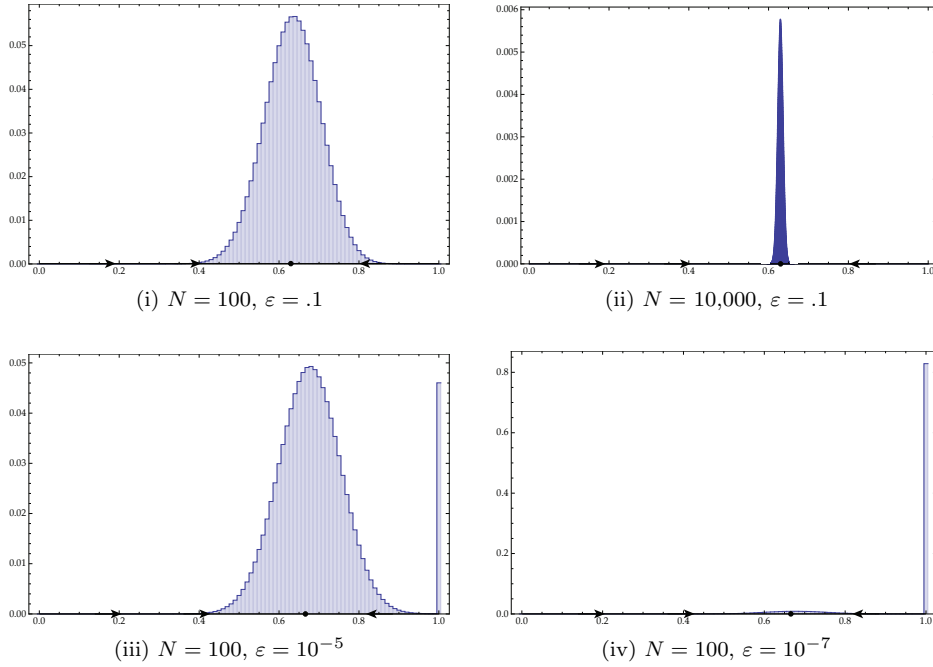$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}.$$

Figure 4: Stationary distribution weights $\mu_\chi^{N,\varepsilon}$ in an anticoordination game under an "imitation with mutation" protocol.

The unique Nash equilibrium of the population game $F(x) = Ax$ is the mixed equilibrium $x^* = (x_0^*, x_1^*) = (\frac{1}{3}, \frac{2}{3})$. To simplify notation in what follows we allow self-matching, but the analysis is virtually identical without it.

Suppose that agents employ the following revision protocol, which combines imitation of successful opponents and mutations:

$$\rho_{ij}^\varepsilon(\pi, x) = x_j \pi_j + \varepsilon.$$

The protocol ρ^ε generates the mean dynamic

$$(19) \qquad \dot{x}_i = V_i^\varepsilon(x) = x_i \hat{F}_i(x) + 2\varepsilon \left(\tfrac{1}{2} - x_i\right),$$

which is the sum of the replicator dynamic and an order ε term that points toward the center of the simplex. When $\varepsilon = 0$, this dynamic is simply the replicator dynamic: the Nash equilibrium $x^* = (\frac{1}{3}, \frac{2}{3})$ attracts solutions from all interior initial conditions, while pure states e_0 and e_1 are unstable rest points. When $\varepsilon > 0$, the two boundary rest points disappear, leaving a globally stable rest point that is near x^*, but slightly closer to the center of the simplex.

Using the formulas from Theorem 5.3, we can compute the stationary distribution $\mu^{N,\varepsilon}$ of the process $\{X_t^{N,\varepsilon}\}$ generated by F and ρ^ε for any fixed values of N and ε. Four instances are presented in Figure 4.

Figure 4(i) presents the stationary distribution when $\varepsilon = .1$ and $N = 100$. This distribution is drawn above the phase diagram of the mean dynamic (19), whose

global attractor is appears at $\hat{\chi} \approx .6296$. The stationary distribution $\mu^{N,\varepsilon}$ has its mode at state $\chi = .64$, but is dispersed rather broadly about this state.

Figure 4(ii) presents the stationary distribution and mean dynamic when $\varepsilon = .1$ and $N = 10{,}000$. Increasing population size moves the mode of the distribution occurs to state $\chi = .6300$, and, more importantly, causes the distribution to exhibit much less dispersion around the modal state. This numerical analysis suggests that in the large population limit, the stationary distribution $\mu^{N,\varepsilon}$ will approach a point mass at $\hat{\chi} \approx .6296$, the global attractor of the relevant mean dynamic.

As the noise level ε approaches zero, the rest point of the mean dynamic approaches the Nash equilibrium $\chi^* = \frac{2}{3}$. Therefore, if after taking N to infinity we take ε to zero, we obtain the double limit

$$\lim_{\varepsilon \to 0} \lim_{N \to \infty} \mu^{N,\varepsilon} = \delta_{\chi^*}, \tag{20}$$

where the limits refer to weak convergence of probability measures, and δ_{χ^*} denotes the point mass at state χ^*.

The remaining pictures illustrate the effects of setting very small mutation rates. When $N = 100$ and $\varepsilon = 10^{-5}$ (Figure 4(iii)) most of the mass in $\mu^{100,\varepsilon}$ falls in a bell-shaped distribution centered at state $\chi = .68$, but a mass of $\mu_1^{100,\varepsilon} = .0460$ sits in isolation at the boundary state $\chi = 1$. When ε is reduced to 10^{-7} (Figure 4(iv)), this boundary state commands a majority of the weight in the disribution ($\mu_1^{100,\varepsilon} = .8286$).

This numerical analysis suggests that when the mutation rate approaches zero, the stationary distribution will approach a point mass at state 1. Increasing the population size does not alter this result, so for the small noise double limit we obtain

$$\lim_{N \to \infty} \lim_{\varepsilon \to 0} \mu^{N,\varepsilon} = \delta_1, \tag{21}$$

where δ_1 denotes the unit point mass at state 1.

Comparing equations (20) and (21), we conclude that the large population double limit and the small noise double limit disagree. ◆

In the preceding example, the large population limits agree with the predictions of the mean dynamic, while the small noise limits do not. Still, the behavior of the latter limits is easy to explain. Starting from any interior state, and from the boundary as well when $\varepsilon > 0$, the expected motion of the process $\{X_t^{N,\varepsilon}\}$ is toward the interior rest point of the mean dynamic V^ε. But when ε is zero, the boundary states 0 and 1 become rest points of V^ε, and are absorbing states of $\{X_t^{N,\varepsilon}\}$; in fact, it is easy to see that they are the only recurrent states of the zero-noise process. Therefore, when $\varepsilon = 0$, $\{X_t^{N,\varepsilon}\}$ reaches either state 0 or state 1 in finite time, and then remains at that state forever.

If instead ε is positive, the boundary states are no longer absorbing, and they are far from any rest point of the mean dynamic. But once the process $\{X_t^{N,\varepsilon}\}$ reaches such a state, it can only depart by way of a mutation. Thus, if we fix the population size N and make ε extremely small, then a journey from an interior state to a boundary state—here a journey against the flow of the mean dynamic—is "more likely" than an escape from a boundary state by way of a single mutation. It follows that in the small noise limit, the stationary distribution must become concentrated on the boundary states regardless of the nature of the mean dynamic. (In fact, it will typically become concentrated on just one of these states.)

As this discussion indicates, the prediction provided by the small noise limit does not become a good approximation of behavior at fixed values of N and ε unless ε is so small that lone mutations are much more rare than excursions from the interior of $\mathcal{X}^N$ to the boundary. In Figures 4(iii) and (iv), which consider a modest population size of $N = 100$, we see that a mutation rate of $\varepsilon = 10^{-5}$ is not small enough to yield agreement with the prediction of the small noise limit, though a mutation rate of $\varepsilon = 10^{-7}$ yields a closer match. With larger population sizes, the relevant mutation rates would be even smaller.

This example suggests that in economic contexts, where the probabilities of "mutations" may not be especially small, the large population limit is more likely to be the relevant one in cases where the predictions of the two limits disagree. In biological contexts, where mutation rates may indeed be quite small, the choice between the limits seems less clear.

6. Asymptotics of the Stationary Distribution and Stochastic Stability

The examples in Section 5 show that even when the underlying game has multiple strict equilibria, the stationary distribution is often concentrated in the vicinity of just one of them if the noise level η is small or the population size N is large. In these cases, the population state so selected provides a unique prediction of infinite horizon play.

In order to obtain clean selection results, we now allow the parameters η and N to approach their limiting values. While each fixed stationary distribution $\mu^{N,\eta}$ has full support on $\mathcal{X}^N$, the limit of a sequence of stationary distributions may converge to a point mass at a single state; thus, taking limits in η and N allows us to obtain exact equilibrium selection results. Moreover while computing a particular stationary distribution requires solving a large collection of linear equalities, the limiting stationary distribution can often be found without explicitly computing any of the stationary distributions along the sequence (see Section 8).

Population states that retain mass in a limiting stationary distribution are said to be *stochastically stable*. There are a number of different definitions of stochastic stability, depending on which limits are taken—just η, just N, η followed by N, or N followed by η—and on what should count as "retaining mass". Taking only the small noise limit, or taking this limit first, emphasizes the rarity of suboptimal play as the key force behind equilibrium selection. Taking only the large population limit, or taking it first, emphasizes the effects of large numbers of conditionally independent decisions in driving equilibrium selection. Since it is not always easy to know which of these forces should be viewed as the primary one, a important goal in stochastic stability analysis is to identify settings in which the small noise and large population limits agree.

Analysis of stochastic stability in games have been carried out under a wide array of assumptions about the form of the underlying game, the nature of the revision protocol, the specification of the evolutionary process, and the limits taken to define stochastic stability—see Section 8 and Sandholm (2009, 2010c) for references. In what follows, we focus on an interesting setting in which all calculations can be carried until their very end, and in which one can obtain precise statements about infinite horizon behavior and about the agreement of the small noise and large population limits.

7. Noisy Best Response Protocols in Two-Strategy Games

Here we consider evolution in two-strategy games under a general class of noisy best response protocols We introduce the notion of the *cost* of a suboptimal choice, which is defined as the rate of decay of the probability of making this choice as the noise level approaches zero. Using this notion, we derive simple formulas that characterize the asymptotics of the stationary distribution under the various limits in η and N, and offer a necessary and sufficient condition for an equilibrium to be uniquely stochastically stable under every noisy best response protocol. This section follows Sandholm (2010a), which builds on earlier work by Binmore and Samuelson (1997), Blume (2003), and Sandholm (2007).

7.1. Noisy Best Response Protocols and their Cost Functions. We consider evolution under *noisy best response protocols*. These protocols can be expressed as

$$(22) \qquad\qquad \rho_{ij}^{\eta}(\pi) = \sigma^{\eta}(\pi_j - \pi_i),$$

for some function $\sigma^{\eta} \colon \mathbb{R} \to (0,1)$: when a current strategy i player receives a revision opportunity, he switches to strategy $j \neq i$ with a probability that only depends on the payoff advantage of strategy j over strategy i. To justify its name, the protocol σ^{η} should recommend optimal strategies with high probability when the noise level is small:

$$\lim_{\eta \to 0} \sigma^{\eta}(a) = \begin{cases} 1 & \text{if } a > 0, \\ 0 & \text{if } a < 0. \end{cases}$$

To place further structure on the probabilities of suboptimal choices, we impose restrictions on the rates at which the probabilities $\sigma^{\eta}(a)$ of choosing a suboptimal strategy approach zero as η approaches zero. To do so, we define the *cost* of switching to a strategy with payoff *disadvantage* $d \in \mathbb{R}$ as

$$(23) \qquad\qquad \kappa(d) = -\lim_{\eta \to 0} \eta \log \sigma^{\eta}(-d).$$

By unpacking this expression, we can write the probability of switching to a strategy with payoff disadvantage d when the noise level is η as

$$\sigma^{\eta}(-d) = \exp\left(-\eta^{-1}(\kappa(d) + o(1))\right),$$

where $o(1)$ represents a term that vanishes as η approaches 0. Thus, $\kappa(d)$ is the exponential rate of decay of the choice probability $\sigma^{\eta}(-d)$ as η^{-1} approaches infinity.

We are now ready to define the class of protocols we will consider.

DEFINITION. We say that the noisy best response protocol (22) is *regular* if
(i)   the limit in (23) exists for all $d \in \mathbb{R}$, with convergence uniform on compact intervals;
(ii) κ is nondecreasing;
(iii) $\kappa(d) = 0$ whenever $d < 0$;
(iv) $\kappa(d) > 0$ whenever $d > 0$.

Conditions (ii)-(iv) impose constraints on the rates of decay of switching probabilities. Condition (ii) requires the rate of decay to be nondecreasing in the payoff disadvantage of the alternative strategy. Condition (iii) requires the switching probability of an agent currently playing the suboptimal strategy to have rate of decay zero; the condition is satisfied when the probability is bounded away from zero, although this is not necessary for the condition to hold. Finally, condition (iv)

requires the probability of switching from the optimal strategy to the suboptimal one to have a positive rate of decay. These conditions are consistent with having either $\kappa(0) > 0$ or $\kappa(0) = 0$: thus, when both strategies earn the same payoff, the probability that a revising agent opts to switch strategies can converge to zero with a positive rate of decay, as in Example 7.1 below, or can be bounded away from zero, as in Examples 7.2 and 7.3.

We now present the three leading examples of noisy best response protocols.

Example 7.1. Best response with mutations. The BRM protocol with noise level $\eta \, (= -(\log \varepsilon)^{-1})$, introduced in Example 5.1, is defined by

$$\sigma^\eta(a) = \begin{cases} 1 - \exp(-\eta^{-1}) & \text{if } a > 0, \\ \exp(-\eta^{-1}) & \text{if } a \leq 0. \end{cases}$$

In this specification, an indifferent agent only switches strategies in the event of a mutation. Since for $d \geq 0$ we have $-\eta \log \sigma^\eta(-d) = 1$, protocol σ^η is regular with cost function

$$\kappa(d) = \begin{cases} 1 & \text{if } d \geq 0, \\ 0 & \text{if } d < 0. \end{cases} \quad \blacklozenge$$

Example 7.2. Logit choice. The *logit choice protocol* with noise level $\eta > 0$, introduced in Examples 3.4 and 5.2, is defined in two-strategy games by

$$\sigma^\eta(a) = \frac{\exp(\eta^{-1}a)}{\exp(\eta^{-1}a) + 1}.$$

For $d \geq 0$, we have that $-\eta \log \sigma^\eta(-d) = d + \eta \log(\exp(-\eta^{-1}d) + 1)$. It follows that σ^η is regular with cost function

$$\kappa(d) = \begin{cases} d & \text{if } d > 0, \\ 0 & \text{if } d \leq 0. \end{cases} \quad \blacklozenge$$

Example 7.3. Probit choice. The logit choice protocol can be derived from a random utility model in which the strategies' payoffs are perturbed by i.i.d., double exponentially distributed random variables (see Hofbauer and Sandholm (2002)). The *probit choice protocol* assumes instead that the payoff perturbations are i.i.d. normal random variables with mean 0 and variance η. Thus

$$\sigma^\eta(a) = \mathbb{P}(\sqrt{\eta}\, Z + a > \sqrt{\eta}\, Z'),$$

where Z and Z' are independent and standard normal. It follows easily that

(24) $$\sigma^\eta(a) = \Phi\left(\tfrac{a}{\sqrt{2\eta}}\right),$$

where Φ is the standard normal distribution function.

A well-known approximation of Φ tells us that when $z < 0$,

(25) $$\Phi(z) = K(z) \exp(\tfrac{-z^2}{2})$$

for some $K(z) \in (\frac{-1}{\sqrt{2\pi}\, z}(1 - \frac{1}{z^2}), \frac{-1}{\sqrt{2\pi}\, z})$. By employing this observation, one can show that σ^η is regular with cost function

$$\kappa(d) = \begin{cases} \tfrac{1}{4}d^2 & \text{if } d > 0, \\ 0 & \text{if } d \leq 0. \end{cases} \quad \blacklozenge$$

7.2. The (Double) Limit Theorem. Our result on the asymptotics of the stationary distribution requires a few additional definitions and assumptions. We suppose that the sequence of two-strategy games $\{F^N\}_{N=N_0}^{\infty}$ converges uniformly to a continuous-population game F, where $F\colon [0,1] \to \mathbb{R}^2$ is a continuous function. We let

$$F_\Delta(\chi) \equiv F_1(\chi) - F_0(\chi)$$

denote the payoff advantage of strategy 1 at state χ in the limit game.

We define the *relative cost function* $\tilde{\kappa}\colon \mathbb{R} \to \mathbb{R}$ by

$$(26) \qquad \tilde{\kappa}(d) = \lim_{\eta \to 0} \left(-\eta \log \sigma^\eta(-d) + \eta \log \sigma^\eta(d) \right) = \kappa(d) - \kappa(-d).$$

Our assumptions on κ imply that $\tilde{\kappa}$ is nondecreasing, sign preserving $(\mathrm{sgn}(\tilde{\kappa}(d)) = \mathrm{sgn}(d))$, and odd $(\tilde{\kappa}(d) = -\tilde{\kappa}(-d))$.

We define the *ordinal potential function* $I\colon [0,1] \to \mathbb{R}$ by

$$(27) \qquad I(\chi) = \int_0^\chi \tilde{\kappa}(F_\Delta(y))\, dy,$$

where the relative cost function $\tilde{\kappa}$ is defined in equation (26). Observe that by marginally adjusting the state χ so as to increase the mass on the optimal strategy, we increase the value of I at rate $\tilde{\kappa}(a)$, where a is the optimal strategy's payoff advantage. Thus, the ordinal potential function combines information about payoff differences with the costs of the associated suboptimal choices.

Finally, we define $\Delta I\colon [0,1] \to (-\infty, 0]$ by

$$(28) \qquad \Delta I(\chi) = I(\chi) - \max_{y \in [0,1]} I(y).$$

Thus, ΔI is obtained from I by shifting its values uniformly, doing so in such a way that the maximum value of ΔI is zero.

Example 7.4. If ρ^η represents best response with mutations (Example 7.1), then the ordinal potential function (27) becomes the *signum potential function*

$$I_{\mathrm{sgn}}(\chi) = \int_0^\chi \mathrm{sgn}\left(F_\Delta(y)\right) dy.$$

This slope of this function at state χ is 1, -1, or 0, according to whether the optimal strategy at χ is strategy 1, strategy 0, or both. ♦

Example 7.5. If ρ^η represents logit choice (Example 7.2), then (27) becomes the (*standard*) *potential function*

$$I_1(\chi) = \int_0^\chi F_\Delta(y)\, dy,$$

whose slope at state χ is just the payoff difference at χ. ♦

Example 7.6. If ρ^η represents probit choice (Example 7.3), then (27) becomes the *quadratic potential function*

$$I_2(\chi) = \int_0^\chi \tfrac{1}{4} \left\langle F_\Delta(y) \right\rangle^2 dy,$$

where $\langle a \rangle^2 = \mathrm{sgn}(a)\, a^2$ is the signed square function. The values of I_2 again depend on payoff differences, but relative to the logit case, larger payoff differences play a more important role. This contrast can be traced to the fact that at small

noise levels, the double exponential distribution has fatter tails than the normal distribution—compare Example 7.3. ♦

Theorem 7.7 shows that whether one takes the small noise limit before the large population limit, or the large population before the small noise limit, the rates of decay of the stationary distribution are captured by the ordinal potential function I. Since the double limits agree, our predictions of infinite horizon behavior under noisy best response rules do not depend on which force drives the equilibrium selection results.

THEOREM 7.7. *The stationary distributions $\mu^{N,\eta}$ satisfy*

(i) $\qquad \lim_{N\to\infty} \lim_{\eta\to 0} \max_{\chi\in\mathcal{X}^N} \left| \frac{\eta}{N} \log \mu_\chi^{N,\eta} - \Delta I(\chi) \right| = 0$ *and*

(ii) $\qquad \lim_{\eta\to 0} \lim_{N\to\infty} \max_{\chi\in\mathcal{X}^N} \left| \frac{\eta}{N} \log \mu_\chi^{N,\eta} - \Delta I(\chi) \right| = 0.$

Theorem 7.7 is proved by manipulating the stationary distribution formula from Theorem 5.3 and applying the dominated convergence theorem.

7.3. Stochastic Stability: Examples and Analysis. Theorem 7.7 describes the rate of decay of the stationary distribution weights as η approaches 0 and N approaches infinity. If the main concern is with the states that are likely to be observed with some frequency over the infinite horizon, then one can focus on states $\chi \in [0,1]$ with $\Delta I(\chi) = 0$, since only neighborhoods of such states receive nonnegligible mass in $\mu^{N,\eta}$ for large N in small η. We therefore call state χ *weakly stochastically stable* if it maximizes the ordinal potential I on the unit interval, and we call state χ *uniquely stochastically stable* if it is the unique maximizer of I on the unit interval. We now investigate in greater detail how a game's payoff function and the revision protocol's cost function interact to determine the stochastically stable states.

Stochastic stability analysis is most interesting when it allows us to select among multiple strict equilibria. For this reason, we focus the analysis to come on coordination games. The two-strategy population game $F : [0,1] \to \mathbb{R}^2$ is a *coordination game* if there is a state $\chi^* \in (0,1)$ such that

$$\mathrm{sgn}(F_\Delta(\chi)) = \mathrm{sgn}(\chi - \chi^*) \ \text{ for all } \chi \neq \chi^*.$$

Any ordinal potential function I for a coordination game is quasiconvex, with local maximizers at each boundary state. Because $I(0) \equiv 0$ by definition, Theorem 7.7 implies the following result.

COROLLARY 7.8. *Suppose that the limit game F is a coordination game. Then state 1 is uniquely stochastically stable in both double limits if $I(1) > 0$, while state 0 is uniquely stochastically stable in both double limits if $I(1) < 0$.*

The next two examples, which revisit two games introduced in the previous chapter, show that the identity of the stochastically stable state may or may not depend on the revision protocol the agents employ.

Example 7.9. Stag Hunt revisited. In Example 5.4, we considered stochastic evolution in the Stag Hunt game

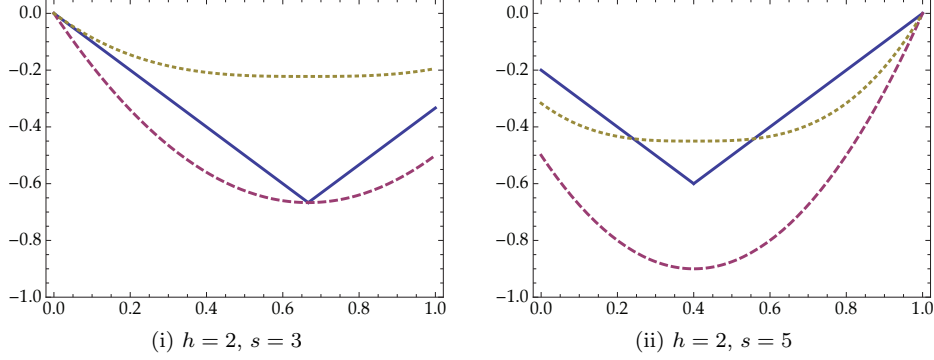$$A = \begin{pmatrix} h & h \\ 0 & s \end{pmatrix},$$

(i) $h = 2$, $s = 3$ (ii) $h = 2$, $s = 5$

Figure 5: The ordinal potentials ΔI_{sgn} (solid), ΔI_1 (dashed), and ΔI_2 (dotted) for Stag Hunt.

where $s > h > 0$. When a continuous population of agents are matched to play this game, their expected payoffs are given by $F_H(\chi) = h$ and $F_S(\chi) = s\chi$, where χ denotes the proportion of agents playing Stag. This coordination game has two pure Nash equilibria, as well as a mixed Nash equilibrium that puts weight $\chi^* = \frac{h}{s}$ on Stag.

The ordinal potentials for the BRM, logit, and probit protocols in this game are

$$I_{\mathrm{sgn}}(\chi) = \left| \chi - \chi^* \right| - \chi^*,$$
$$I_1(x) = \tfrac{s}{2}\chi^2 - h\chi, \text{ and}$$
$$I_2(x) = \begin{cases} -\frac{s^2}{12}\chi^3 + \frac{hs}{4}\chi^2 - \frac{h^2}{4}\chi & \text{if } \chi \le \chi^*, \\ \frac{s^2}{12}\chi^3 - \frac{hs}{4}\chi^2 + \frac{h^2}{4}\chi - \frac{h^3}{6s} & \text{if } \chi > \chi^*. \end{cases}$$

Figure 5 presents the normalized functions ΔI_{sgn}, ΔI_1, and ΔI_2 for two specifications of payoffs: $h = 2$ and $s = 3$ (in (i)), and $h = 2$ and $s = 5$ (in (ii)). For any choices of $s > h > 0$, ΔI is symmetric about its minimizer, the mixed Nash equilibrium $\chi^* = \frac{h}{s}$. As a result, the three protocols always agree about equilibrium selection: the all-Hare equilibrium is uniquely stochastically stable when $\chi^* > \frac{1}{2}$ (or, equivalently, when $2h > s$), while the all-Stag equilibrium is uniquely stochastically stable when the reverse inequality holds. ◆

Example 7.10. Nonlinear Stag Hunt revisited. In Example 5.5, we introduced the nonlinear Stag Hunt game with payoff functions $F_H(\chi) = h$ and $F_S(\chi) = s\chi^2$, with χ again representing the proportion of agents playing Stag. This game has two pure Nash equilibria and a mixed equilibrium at $\chi^* = \sqrt{h/s}$. The payoffs and mixed equilibria for $h = 2$ and various choices of s are graphed in Figure 6.

The ordinal potentials for the BRM, logit, and probit models are given by

$$I_{\mathrm{sgn}}(\chi) = \left| \chi - \chi^* \right| - \chi^*,$$
$$I_1(x) = \tfrac{s}{3}\chi^3 - h\chi, \text{ and}$$
$$I_2(x) = \begin{cases} -\frac{s^2}{20}x^5 + \frac{hs}{6}x^3 - \frac{h^2}{4}x & \text{if } \chi \le \chi^*, \\ \frac{s^2}{20}x^5 - \frac{hs}{6}x^3 + \frac{h^2}{4}x - \frac{4h^2\chi^*}{15} & \text{if } \chi > \chi^*. \end{cases}$$
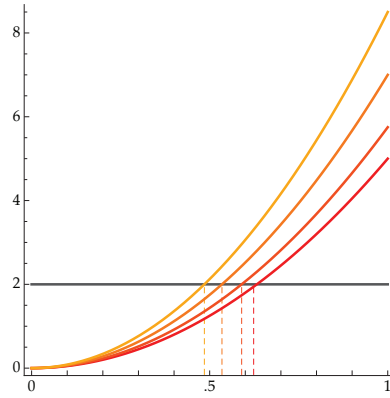
Figure 6: Payoffs and mixed equilibria in Nonlinear Stag Hunt when $h = 2$ and $s = 5, 5.75, 7$, and 8.5.

(i) $h = 2$, $s = 5$

(ii) $h = 2$, $s = 5.75$

(iii) $h = 2$, $s = 7$
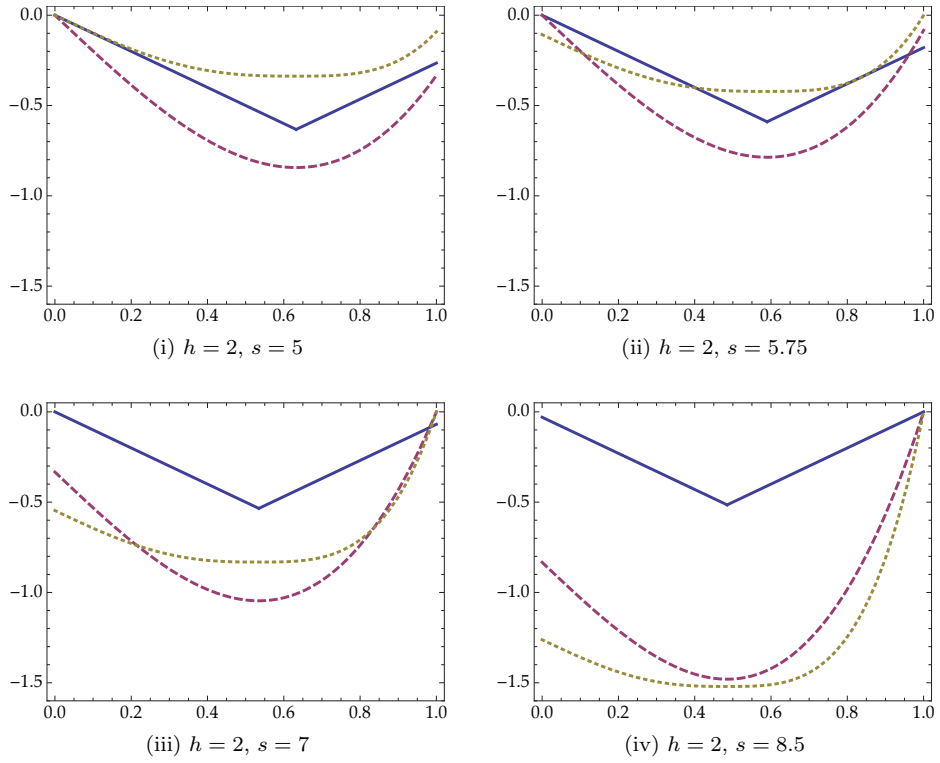
(iv) $h = 2$, $s = 8.5$

Figure 7: The ordinal potentials ΔI_{sgn} (solid), ΔI_1 (dashed), and ΔI_2 (dotted) for Nonlinear Stag Hunt.

Figure 7 presents the functions ΔI_{sgn}, ΔI_1, and ΔI_2 for $h = 2$ and for various choices of s.

When s is at its lowest level of 5, coordination on Stag is at its least appealing. Since $\chi^* = \sqrt{2/5} \approx .6325$, the basin of attraction of the all-Hare equilibrium is considerably larger than that of the all-Stag equilibrium. Figure 7(i) illustrates that coordination on Hare is stochastically stable under all three protocols.

If we make coordination on Stag somewhat more attractive by increasing s to 5.75, the mixed equilibrium becomes $\chi^* = \sqrt{2/5.75} \approx .5898$. The all-Hare equilibrium remains stochastically stable under the BRM and logit rules, but all-Stag becomes stochastically stable under the probit rule (Figure 7(ii)).

Increasing s further to 7 shifts the mixed equilibrium closer to the midpoint of the unit interval ($\chi^* = \sqrt{2/7} \approx .5345$). The BRM rule continues to select all-Hare, while the probit and logit rules both select all-Stag (Figure 7(iii)).

Finally, when $s = 8.5$, the all-Stag equilibrium has the larger basin of attraction ($\chi^* = \sqrt{2/8.5} \approx .4851$). At this point, coordination on Stag becomes attractive enough that all three protocols select the all-Stag equilibrium (Figure 7(iv)).

Why as we increase the value of s does the transition to selecting all-Stag occur first for the probit rule, then for the logit rule, and finally for the BRM rule? Examining Figure 6, we see that increasing s not only shifts the mixed Nash equilibrium to the left, but also markedly increases the payoff advantage of Stag at states where it is optimal. Since the cost function of the probit rule is the most sensitive to payoff differences, its equilibrium selection changes at the lowest level of s. The next selection to change is that of the (moderately sensitive) logit rule, and the last is the selection of the (insensitive) BRM rule. ◆

7.4. Risk Dominance, Stochastic Dominance, and Stochastic Stability. Building on these examples, we now seek general conditions on payoffs that ensure stochastic stability under all noisy best response protocols.

Example 7.9 showed that in the Stag Hunt game with linear payoffs, the noisy best response rules we considered always selected the equilibrium with the larger basin of attraction. The reason for this is easy to explain. Linearity of payoffs, along with the fact that the relative cost function $\tilde{\kappa}$ is sign-preserving and odd (see equation (26)), implies that the ordinal potential function I is symmetric about the mixed equilibrium χ^*, where it attains its minimum value. If, for example, χ^* is less than $\frac{1}{2}$, so that pure equilibrium 1 has the larger basin of attraction, then $I(1)$ exceeds $I(0)$, implying that state 1 is uniquely stochastically stable. Similarly, if χ^* exceeds $\frac{1}{2}$, then $I(0)$ exceeds $I(1)$, and state 0 is uniquely stochastically stable.

With this motivation, we call strategy i *strictly risk dominant* in the two-strategy coordination game F if the set of states where it is the unique best response is larger than the corresponding set for strategy $j \neq i$. Thus, if F has mixed equilibrium $\chi^* \in (0, 1)$, then strategy 0 is strictly risk dominant if $\chi^* > \frac{1}{2}$, and strategy 1 is strictly risk dominant if $\chi^* < \frac{1}{2}$. If the relevant inequality holds weakly in either case, we call the strategy in question *weakly risk dominant*.

The foregoing arguments yield the following result, in which we denote by e_i the state at which all agents play strategy i.

COROLLARY 7.11. *Suppose that the limit game F is a coordination game with linear payoffs. Then*

(i) *State e_i is weakly stochastically stable under every noisy best response protocol if and only if strategy i is weakly risk dominant in F.*

(*ii*)   *If strategy i is strictly risk dominant in F, then state e_i is uniquely stochastically stable under every noisy best response protocol.*

Example 7.10 shows that once we turn to games with nonlinear payoffs, risk dominance only characterizes stochastic stability under the BRM rule. In any coordination game with mixed equilibrium χ^*, the ordinal potential function for the BRM rule is $I_{\mathrm{sgn}}(\chi) = |\chi - \chi^*| - \chi^*$. This function is minimized at χ^*, and increases at a unit rate as one moves away from χ^* in either direction, reflecting the fact that under the BRM rule, the probability of a suboptimal choice is independent of its payoff consequences. Clearly, whether $I_{\mathrm{sgn}}(1)$ is greater than $I_{\mathrm{sgn}}(0)$ depends only on whether χ^* is less than $\frac{1}{2}$. We therefore have

COROLLARY 7.12. *Suppose that the limit game F is a coordination game and that σ^η is the BRM rule. Then*
   (*i*)   *State e_i is weakly stochastically stable if and only if strategy i is weakly risk dominant in F.*
   (*ii*)   *If strategy i is strictly risk dominant in F, then state e_i is uniquely stochastically stable.*

Once one moves beyond the BRM rule and linear payoffs, risk dominance is no longer a necessary or sufficient condition for stochastic stability. In what follows, we introduce a natural refinement of risk dominance that serves this role.

To work toward our new definition, let us first observe that any function on the unit interval $[0, 1]$ can be viewed as a random variable by regarding the interval as a sample space endowed with Lebesgue measure λ. With this interpretation in mind, we define the *advantage distribution* of strategy i to be the cumulative distribution function of the payoff advantage of strategy i over the alternative strategy $j \neq i$:

$$G_i(a) = \lambda(\{\chi \in [0, 1] : F_i(\chi) - F_j(\chi) \leq a\}).$$

We let $\bar{G}_i$ denote the corresponding decumulative distribution function:

$$\bar{G}_i(a) = \lambda(\{\chi \in [0, 1] : F_i(\chi) - F_j(\chi) > a\}) = 1 - G_i(a).$$

In words, $\bar{G}_i(a)$ is the measure of the set of states at which the payoff to strategy i exceeds the payoff to strategy j by more than a.

It is easy to restate the definition of risk dominance in terms of the advantage distribution.

OBSERVATION 7.13. *Let F be a coordination game. Then strategy i is weakly risk dominant if and only if $\bar{G}_i(0) \geq \bar{G}_j(0)$, and strategy i is strictly risk dominant if and only if $\bar{G}_i(0) > \bar{G}_j(0)$.*

To obtain our refinement of risk dominance, we require not only that strategy i be optimal at a larger set of states than strategy j, but also that strategy i have a payoff advantage of at least a at a larger set of states than strategy j for every $a \geq 0$. More precisely, we say that strategy i is *weakly stochastically dominant* in the coordination game F if $\bar{G}_i(a) \geq \bar{G}_j(a)$ for all $a \geq 0$. If in addition $\bar{G}_i(0) > \bar{G}_j(0)$, we say that strategy i is *strictly stochastically dominant*. The notion of stochastic dominance for strategies proposed here is obtained by applying the usual definition of stochastic dominance from utility theory (see Border (2001)) to the strategies' advantage distributions.

Theorem 7.14 shows that stochastic dominance is both sufficient and necessary to ensure stochastic stability under every noisy best response rule.

THEOREM 7.14. *Suppose that the limit game F is a coordination game. Then*
(i)   *State e_i is weakly stochastically stable under every noisy best response protocol if and only if strategy i is weakly stochastically dominant in F.*
(ii)   *If strategy i is strictly stochastically dominant in F, then state e_i is uniquely stochastically stable under every noisy best response protocol.*

The idea behind Theorem 7.14 is simple. The definitions of I, $\tilde{\kappa}$, κ, F_Δ, and G_i imply that

$$
(29) \qquad I(1) = \int_0^1 \tilde{\kappa}(F_\Delta(y)) \, \mathrm{d}y
$$

$$
= \int_0^1 \kappa(F_1(y) - F_0(y)) \, \mathrm{d}y - \int_0^1 \kappa(F_0(y) - F_1(y)) \, \mathrm{d}y
$$

$$
= \int_{-\infty}^\infty \kappa(a) \, \mathrm{d}G_1(a) - \int_{-\infty}^\infty \kappa(a) \, \mathrm{d}G_0(a).
$$

As we have seen, whether state e_1 or state e_0 is stochastically stable depends on whether $I(1)$ is greater than or less than $I(0) = 0$. This in turn depends on whether the value of the first integral in the final line of (29) exceeds the value of the second integral. Once we recall that the cost function κ is monotone, Theorem 7.14 reduces to a variation on the standard characterization of first-order stochastic dominance: namely, that distribution G_1 stochastically dominates distribution G_0 if and only if $\int \kappa \, \mathrm{d}G_1 \geq \int \kappa \, \mathrm{d}G_0$ for every nondecreasing function κ.

8. Further Developments

The analyses in the previous sections have focused on evolution in two-strategy games, mostly under noisy best response protocols. Two-strategy games have the great advantage of generating birth-and-death processes. Because such processes are reversible, their stationary distributions can be computed explicitly, greatly simplifying the analysis. Other work in stochastic evolutionary game theory focusing on birth-and-death chain models includes Binmore and Samuelson (1997), Maruta (2002), Blume (2003), and Sandholm (2011). The only many-strategy evolutionary game environments known to generate reversible processes are potential games (Monderer and Shapley (1996); Sandholm (2001)), with agents using either the standard (Example 5.2) or imitative versions of the logit choice rule; see Blume (1993, 1997) and Sandholm (2011) for analyses of these models.

Once one moves beyond reversible settings, obtaining exact formulas for the stationary distribution is generally impossible, and one must attempt to determine the stochastically stable states by other means. In general, the available techniques for doing so are descendants of the analyses of sample path large deviations due to Freidlin and Wentzell (1998), and introduced to evolutionary game theory by Kandori et al. (1993) and Young (1993).

One portion of the literature considers *small noise limits*, determining which states retain mass in the stationary distribution as the amount of noise in agents' decisions vanishes. The advantage of this approach is that the set of population states stays fixed and finite. This makes it possible to use the ideas of Freidlin and Wentzell (1998) with few technical complications, but also without the computational advantages that a continuous state space can provide. Many of the analyses of small noise limits focus on the best response with mutations model

(Example 5.1); see Kandori et al. (1993), Young (1993, 1998), Kandori and Rob (1995, 1998), Ellison (2000), and Beggs (2005). Analyses of other important models include Myatt and Wallace (2003), Fudenberg and Imhof (2006, 2008), Dokumacı and Sandholm (2011), and Staudigl (2011).

Alternatively, one can consider *large population limits*, examining the behavior of the stationary distribution as the population size approaches infinity. Here, as one increases the population size, the set of population states becomes an increasingly fine grid in the simplex X. While this introduces some technical challenges, it also allows one to use methods from optimal control theory in the analysis of sample path large deviations. The use of large population limits in stochastic evolutionary models was first proposed by Binmore and Samuelson (1997) and Blume (2003) in two-strategy settings. Analyses set in more general environments include Benaïm and Weibull (2003) and Benaïm and Sandholm (2011), both of which build on results in Benaïm (1998). The analysis of infinite-horizon behavior in the large population limit is still at an early stage of development, and so offers a promising avenue for future research.

References

Beckmann, M., McGuire, C. B., and Winsten, C. B. (1956). *Studies in the Economics of Transportation*. Yale University Press, New Haven.

Beggs, A. W. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory*, 122:1–36.

Benaïm, M. (1998). Recursive algorithms, urn processes, and the chaining number of chain recurrent sets. *Ergodic Theory and Dynamical Systems*, 18:53–87.

Benaïm, M. and Sandholm, W. H. (2011). Large deviations, reversibility, and equilibrium selection under evolutionary game dynamics. Unpublished manuscript, Université de Neuchâtel and University of Wisconsin.

Benaïm, M. and Weibull, J. W. (2003). Deterministic approximation of stochastic evolution in games. *Econometrica*, 71:873–903.

Binmore, K. and Samuelson, L. (1997). Muddling through: Noisy equilibrium selection. *Journal of Economic Theory*, 74:235–265.

Björnerstedt, J. and Weibull, J. W. (1996). Nash equilibrium and evolution by imitation. In Arrow, K. J. et al., editors, *The Rational Foundations of Economic Behavior*, pages 155–181. St. Martin's Press, New York.

Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5:387–424.

Blume, L. E. (1997). Population games. In Arthur, W. B., Durlauf, S. N., and Lane, D. A., editors, *The Economy as an Evolving Complex System II*, pages 425–460. Addison-Wesley, Reading, MA.

Blume, L. E. (2003). How noise matters. *Games and Economic Behavior*, 44:251–271.

Border, K. C. (2001). Comparing probability distributions. Unpublished manuscript, Caltech.

Brown, G. W. and von Neumann, J. (1950). Solutions of games by differential equations. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games I*, volume 24 of *Annals of Mathematics Studies*, pages 73–79. Princeton University Press, Princeton.

Dokumacı E. and Sandholm, W. H. (2011). Large deviations and multinomial probit choice. *Journal of Economic Theory*, forthcoming.

Ellison, G. (2000). Basins of attraction, long run equilibria, and the speed of step-by-step evolution. *Review of Economic Studies*, 67:17–45.

Freidlin, M. I. and Wentzell, A. D. (1998). *Random Perturbations of Dynamical Systems.* Springer, New York, second edition.

Fudenberg, D. and Imhof, L. A. (2006). Imitation processes with small mutations. *Journal of Economic Theory*, 131:251–262.

Fudenberg, D. and Imhof, L. A. (2008). Monotone imitation dynamics in large populations. *Journal of Economic Theory*, 140:229–245.

Fudenberg, D. and Levine, D. K. (1998). *The Theory of Learning in Games.* MIT Press, Cambridge.

Gilboa, I. and Matsui, A. (1991). Social stability and equilibrium. *Econometrica*, 59:859–867.

Helbing, D. (1992). A mathematical model for behavioral changes by pair interactions. In Haag, G., Mueller, U., and Troitzsch, K. G., editors, *Economic Evolution and Demographic Change: Formal Models in Social Sciences*, pages 330–348. Springer, Berlin.

Hofbauer, J. (1995). Imitation dynamics for games. Unpublished manuscript, University of Vienna.

Hofbauer, J. and Sandholm, W. H. (2002). On the global convergence of stochastic fictitious play. *Econometrica*, 70:2265–2294.

Hofbauer, J. and Sigmund, K. (1988). *Theory of Evolution and Dynamical Systems.* Cambridge University Press, Cambridge.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge.

Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society (New Series)*, 40:479–519.

Kandori, M., Mailath, G. J., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, 61:29–56.

Kandori, M. and Rob, R. (1995). Evolution of equilibria in the long run: A general theory and applications. *Journal of Economic Theory*, 65:383–414.

Kandori, M. and Rob, R. (1998). Bandwagon effects and long run technology choice. *Games and Economic Behavior*, 22:84–120.

Kurtz, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7:49–58.

Maruta, T. (2002). Binary games with state dependent stochastic choice. *Journal of Economic Theory*, 103:351–376.

Maynard Smith, J. (1982). *Evolution and the Theory of Games.* Cambridge University Press, Cambridge.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14:124–143.

Moran, P. A. P. (1962). *The Statistical Processes of Evolutionary Theory.* Clarendon Press, Oxford.

Myatt, D. P. and Wallace, C. C. (2003). A multinomial probit model of stochastic evolution. *Journal of Economic Theory*, 113:286–301.

Norris, J. R. (1997). *Markov Chains.* Cambridge University Press, Cambridge.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life.* Belknap/Harvard, Cambridge.

Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428:646–650.

Sandholm, W. H. (2001). Potential games with continuous player sets. *Journal of Economic Theory*, 97:81–108.

Sandholm, W. H. (2003). Evolution and equilibrium under inexact information. *Games and Economic Behavior*, 44:343–378.

Sandholm, W. H. (2007). Simple formulas for stationary distributions and stochastically stable states. *Games and Economic Behavior*, 59:154–162.

Sandholm, W. H. (2009). Evolutionary game theory. In Meyers, R. A., editor, *Encyclopedia of Complexity and Systems Science*, pages 3176–3205. Springer, Heidelberg.

Sandholm, W. H. (2010a). Orders of limits for stationary distributions, stochastic dominance, and stochastic stability. *Theoretical Economics*, 5:1–26.

Sandholm, W. H. (2010b). Pairwise comparison dynamics and evolutionary foundations for Nash equilibrium. *Games*, 1:3–17.

Sandholm, W. H. (2010c). *Population Games and Evolutionary Dynamics.* MIT Press, Cambridge.

Sandholm, W. H. (2011). Stochastic imitative game dynamics with committed agents. Unpublished manuscript, University of Wisconsin.

Schlag, K. H. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78:130–156.

Smith, M. J. (1984). The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov. *Transportation Science*, 18:245–252.

Staudigl, M. (2011). Stochastic stability in binary choice coordination games. Unpublished manuscript, European University Institute.

Taylor, P. D. and Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156.

Traulsen, A. and Hauert, C. (2009). Stochastic evolutionary game dynamics. In Schuster, H. G., editor, *Reviews of Nonlinear Dynamics and Complexity*, volume 2, pages 25–61. Wiley, New York.

Weibull, J. W. (1995). *Evolutionary Game Theory.* MIT Press, Cambridge.

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

Young, H. P. (1998). *Individual Strategy and Social Structure.* Princeton University Press, Princeton.

DEPARTMENT OF ECONOMICS, UNIVERSITY OF WISCONSIN, 1180 OBSERVATORY DRIVE, MADISON, WI 53706, USA.

E-mail address: `whs@ssc.wisc.edu`