Finite-population evolution with rare mutations in asymmetric games

Carl Veller†‡ Laura K. Hayward§

March 20, 2015

Abstract

We model evolution according to an asymmetric game as occurring in multiple finite populations, one for each role in the game, and study the effect of subjecting individuals to stochastic strategy mutations. We show that, when these mutations occur sufficiently infrequently, the dynamics over all population states simplify to an ergodic Markov chain over just the pure population states (where each population is monomorphic). This makes calculation of the stationary distribution computationally feasible. The transition probabilities of this embedded Markov chain involve fixation probabilities of mutants in single populations. The asymmetry of the underlying game leads to fixation probabilities that are derived from frequency-independent selection, in contrast to the analogous single-population symmetric-game case (Fudenberg and Imhof, 2006). This frequency-independence is useful in that it allows us to employ results from the population genetics literature to calculate the stationary distribution of the evolutionary process, giving sharper, and sometimes even analytic, results. We demonstrate the utility of this approach by applying it to a battle-of-the-sexes game, a Crawford-Sobel signalling game, and the beer-quiche game of Cho and Kreps (1987).

1 Introduction

In asymmetric games, Nash equilibria are often not strict. This is especially true for asymmetric multi-stage games, where alternative strategies that induce the same path of play, and prescribe the same actions on that path but different actions off it, are payoff equivalent. Because of this, the classical evolutionary approach to equilibrium selection, involving deterministic evolutionary dynamics in infinite populations, and centered on the concept of evolutionary stability, has little to offer for such games: for a Nash equilibrium of an asymmetric game to be evolutionarily stable, it must be strict (Samuelson and Zhang, 1992).
This is a major shortcoming, because the class of asymmetric games includes many games of major importance to economists and biologists—for example, signalling games (Spence, 1973; Crawford and Sobel, 1982; Grafen, 1990; Veller et al., 2015), games of entry and entry-deterrence (Salop, 1979; Milgrom and Roberts, 1982; Maynard Smith and Parker, 1976), and games of ‘time consistency’ and commitment (Kydland and Prescott, 1977). Evolutionary game theory provides a natural method by which to model learning processes (Fudenberg and Levine, 1998) and biological evolution (Maynard Smith, 1982), and provides a ‘rationality-light’ approach to equilibrium selection in games (Samuelson, 1998). These are things we would surely like to be able to do for contexts that are well modelled by asymmetric games.

In developing a more fruitful evolutionary approach to equilibrium selection in asymmetric games, and asymmetric multi-stage games in particular, it is instructive to consider precisely why the deterministic, infinite-population approach often fails in these games. When none of the Nash equilibria of a game is strict, populations that play in accordance with them can be invaded by ‘neutral mutants’ that achieve the same payoff as the incumbents. Therefore, none of these equilibria is evolutionarily stable, and this method of equilibrium selection fails. Moreover, we typically think of such ‘neutral’ invasions as random processes, where strategies that do as well as each other fluctuate in their relative representations in the population. Deterministic dynamics cannot model the stochasticity that governs this ‘neutral drift’. (The replicator dynamics, for example, remain stationary after the invasion of a small measure of neutral mutants.)

A more natural approach to modelling neutral drift, and thus evolution in games that do not admit strict Nash equilibria, is to treat evolution as a stochastic process occurring in a finite population (or populations). This approach also allows us to more naturally model stochastic mutations (thought of as errors or experimentation under the learning interpretation of evolutionary dynamics). Indeed, it is for these reasons that the theory of finite-population evolution has been extensively studied in the population genetics literature (Crow and Kimura, 1970; Ewens, 2004), although a major barrier to employing results from this literature in evolutionary game theory is that selection in population genetics is almost always treated as frequency-independent.

Finite-population dynamics also allow us to dispense with the stability concepts for equilibrium selection that infinite-population dynamics rely on, and which so often fail for asymmetric games of interest. Instead, when finite-population dynamics constitute an ergodic Markov chain over all population states (e.g., when mutation rates are positive), we can focus on the masses of the various population states in the stationary distribution of this Markov chain (Foster and Young, 1990; Kandori et al., 1993). This allows for equilibrium selection even in games that do not admit an evolutionarily stable outcome, and in any case, offers a useful robustness check for other equilibrium selection concepts.

This paper analyzes finite-population evolutionary dynamics in asymmetric games. Since asymmetric games are characterized by the existence of multiple ‘roles’, the simplest evolutionary model involves multiple interacting populations, one for each role. Selection
within each population acts (stochastically) according to the relative fitnesses of strategies employed in that population. These fitnesses derive from the expected payoff each strategy receives, according to the underlying game, from interacting with the other populations. We assume positive mutation rates, entailing a positive probability that any given individual in any period changes her strategy, with all alternative choices from her strategy space equally likely (though this last assumption can easily be relaxed—see Section 5). In this setup the overall population dynamics constitute an ergodic Markov chain. We focus on its stationary distribution.

One potential drawback is that the state space of this Markov chain, viz., all population states, is typically very large, making calculation of the stationary distribution computationally infeasible. One possible workaround is to analyze the limit of the stationary distributions as mutation rates become small across all populations; this approach has previously been studied for symmetric games, where evolution proceeds in a single population with or without population structure (Fudenberg and Imhof, 2006).

Here, we prove that, in the case of multiple populations interacting according to an asymmetric game, as the mutation rates approach zero in all populations, the stationary distribution of the evolutionary Markov chain over all population states approaches that of a Markov chain over a much smaller state space, the pure population states (in which every population is monomorphic). In this embedded Markov chain, the one-step transition probabilities between pure states that differ by more than one population’s strategy are all zero. The transition probability between two pure states that differ by only one population’s strategy is proportional to the probability of a mutant appearing in the appropriate population in the antecedent state multiplied by the probability that the mutant, having appeared, subsequently takes over that population (‘fixes’) under the evolutionary process without mutations.

Since the state space of this embedded Markov chain is significantly smaller than the set of all population states, calculation of the invariant distribution of the embedded Markov chain will typically be computationally feasible. Moreover, unlike in the single-population case studied in Fudenberg and Imhof (2006), the fixation probabilities that are used to construct the transition probability matrix of this embedded Markov chain derive from frequency-independent selection. This is a result of the asymmetry of the underlying game: the fitness difference between the incumbent and mutant strategies in the focal population depend only on the population states of the other populations, and these populations remain monomorphic under the no-mutation process. This is a significant advantage, because fixation probabilities for most well-known evolutionary and imitation processes, when selection is frequency-dependent, either do not exist in closed form, or are intractable when they do (Nowak, 2006). On the other hand, when selection is frequency-independent, there often exist relatively simple closed form representations of the fixation probabilities of well-known processes (exact or approximate), and we may thus employ results from population genetics and other literatures to derive sharper, and sometimes even analytical, results for many asymmetric games of interest. This allows for
powerful evolutionary equilibrium selection in these games, and also allows us to study the
influence of various parameters of the learning or evolutionary process (e.g., the rate of
mutation/experimentation, the strength of selection/imitation, the size of the populations)
on its outcome. This in turn allows us to pose, and in a transparent way solve, questions
such as: ‘For a particular game, does a higher rate of experimentation in one of the
populations improve that population’s long run success?’, ‘If one population is much larger
than another, which fares better in the long run?’, etc.

We demonstrate the utility of our result by applying it to a simple ‘battle of the sexes’
game, a discrete version of the signalling game from Crawford and Sobel (1982), and the
‘beer-quiche’ game of Cho and Kreps (1987).

2 Evolution with mutations in multiple finite populations

Asymmetric games are characterized by the existence of multiple ‘roles’ (‘Player 1’, ‘Player
2’, etc.). In the evolutionary approach, the simplest way to incorporate multiple roles is
to model evolution as occurring in multiple interacting populations.\footnote{In Section 5, we discuss how our results relate to the alternative modelling case of a single population, in which each generation, each member draws a role from some distribution.}

Suppose that we have an underlying game \( \Gamma \) with roles \( i = 1, \ldots, I \), each role associated
with a strategy set \( S_i \), and the payoff to a player in role \( i \) when play is \( (s_1, \ldots, s_I) \in \prod_{i=1}^I S_i \)
given by \( \pi_i(s_1, \ldots, s_I) \in \mathbb{R} \).

We assume the existence of \( I \) populations, one for each role, with the size of each
population \( i \) constant through time at \( N_i \in \mathbb{N} \). The overall population state at a given
time is defined as the \( I \)-tuple of strategy frequencies in the respective populations at that
time: \( p^t \in \prod_{i=1}^I \Delta_{|S_i|} \), where \( \Delta^n \) is the unit simplex in \( \mathbb{R}^n \).\footnote{Since the populations are finite, \( p^t \) is in fact confined to a finite subset of this space.} We shall be interested in the
evolution of this population state over time.

Evolution proceeds as a stochastic process in discrete time. Each generation, each
member of each population receives the expected value of interacting, according to \( \Gamma \),
with a group comprising one member from each other population, randomly chosen, and
with each group equally likely. (The use of expected payoffs, rather than true payoffs
received from single random interactions, is for the sake of tractability.) If \( \pi^t_j(s^k_j) \) denotes
the proportion of members of population \( j \) that are playing strategy \( s^k_j \in S_j \) at time \( t \),
then, for example, the expected payoff to a member of population 1 who employs strategy
\( s^1_1 \in S_1 \) in period \( t \) is

\[
\mathbb{E}\pi_1(s^1_1|p^t) = \mathbb{E}\pi_1(s^1_1|p^t_{-1}) = \sum_{k_2=1}^{|S_2|} \cdots \sum_{k_I=1}^{|S_I|} \pi^t_2(s^k_2) \cdots \pi^t_I(s^k_I) \pi_1(s^1_1, s^k_2, \ldots, s^k_I).
\]

Here, \( p^t_{-1} \) denotes the population states in all populations other than population 1, and
signifies that the expected payoff to a strategy in population 1 depends only on the strategy
frequencies in the other populations 2, . . . , I. This is a consequence of the asymmetry of
the underlying game.

These expected payoffs in each population \( i \) are then translated to non-negative fitnesses \( f_i(s_i^t|p_{-i}) \) according to some positive monotonic transformation (possibly different for each population).\(^3\) In the case of no mutations, the fitnesses within each population can be used to update that population to its next-period state according to an evolutionary, or imitation, dynamic, usually following the general Darwinian, or ‘monotonicity’, principle that strategies with high fitness increase in proportion relative to those with low fitness.

Some notation: let \( \mathcal{P}_i \) denote the (finite) set of all possible population states for population \( i \), let \( \mathcal{P} = \prod_{i=1}^I \mathcal{P}_i \) denote the set of all possible overall population states, and let \( \mathcal{P}_{-i} \) denote the set of all possible population states for populations other than \( i \). The set of ‘pure’ states for population \( i \), \( \mathcal{P}_i^\text{pure} \), comprises all states in \( \mathcal{P}_i \) where every member of population \( i \) is playing the same strategy (in which case we say that population \( i \) is ‘monomorphic’). Abusing notation a little, we label such states by the strategy that all members are playing, i.e., \( \mathcal{P}_i^\text{pure} = S_i \). Finally, the set of overall pure states, \( \mathcal{P}^\text{pure} = \prod_{i=1}^I \mathcal{P}_i^\text{pure} \), is the set of overall population states in which every population is pure.

The evolutionary process with no mutations in each population \( i \) is a stochastic process \( \{X_i^t(t), t = 0, 1, \ldots \} \), with state space \( \mathcal{P}_i \), and transition probabilities \( T_i^0(p_i, p'_i|p_{-i}) \) for \( p = \langle p_i, p_{-i} \rangle \in \mathcal{P}, p'_i \in \mathcal{P}_i \). The transition probabilities depend on the population state \( p_{-i} \) because this determines fitnesses within population \( i \).

For each population \( i \), we require two basic assumptions of this no-mutation evolutionary process defined by \( T_i^0(p_i, p'_i|p_{-i}) \):

Assumption 1. For any \( p_{-i} \), if in some period a strategy in population \( i \) is absent, then it is absent in the next period. Formally, for all \( \langle p_i, p_{-i} \rangle \in \mathcal{P}, p'_i \in \mathcal{P}_i, \) and \( s_i \in S_i \), if \( p_i(s_i) = 0 \) and \( T(p_i, p'_i|p_{-i}) > 0 \), then \( p'_i(s_i) = 0 \).

Assumption 2. For any \( \langle p_i, p_{-i} \rangle \in \mathcal{P}, \) and for each \( s_i \in S_i \) such that \( 0 < p_i(s_i) < 1 \), there exists \( p'_i \in \mathcal{P}_i \) such that \( p'_i(s_i) > p_i(s_i) \) and \( T_i^0(p_i, p'_i|p_{-i}) > 0 \). That is, no matter the population state of other populations, any strategy currently played in \( i \), unless it is played by all members of \( i \), has positive probability of having increased representation next period.

Loosely, Assumption 1 ensures that the pure states for population \( i \) are absorbing. Assumption 2 ensures that, no matter what the population states of the other populations, non-pure states in population \( i \) are transient.

These assumptions are satisfied by many of the standard finite-population stochastic processes studied in the evolutionary game theory and population genetics literatures.

\(^3\)Popular choices in the evolutionary game theory literature include exponential fitness, \( f_i(X_{-i}) = \exp(\eta_i\bar{x}_{-i}) \), and linear fitness, \( f_i(X_{-i}) = 1 + \eta_i\bar{x}_{-i} \); in each case, the parameter \( \eta_i > 0 \) mediates the strength of selection, i.e., the sensitivity of fitness to changes in expected payoff.
when mutations are disallowed, selection is only finitely strong, and fitnesses are positive and finite. Such processes include stochastic models of imitation learning (Fudenberg and Levine, 1998), the Moran process (Moran, 1958), the Wright-Fisher process (Fisher, 1930; Wright, 1931), and the Cannings exchangeable model (Cannings, 1974). Well-known processes that are excluded are the best-response dynamics and fictitious play (Fudenberg and Levine, 1998).

We make the further assumption that the evolutionary processes occur independently within each population, in the sense that, although the probability that population $i$ transitions from $p_i$ to $p'_i$ between periods $t$ and $t + 1$ depends on the period-$t$ population states of the other populations, the transitions that these other populations make between periods $t$ and $t + 1$ do not influence the transition in population $i$. This is similar to the assumption that expected, rather than realized, payoffs are relevant for fitnesses, in the sense that it too is an abstraction from the true, random, matching of players in a given period. Like the expected payoffs assumption, it is made for tractability.

Under this assumption, the no-mutation processes $\{T_i^0\}_{i=1}^I$ aggregate to an overall no-mutation Markov process $T^0$ over the state space $\mathcal{P}$, where for $p = \langle p_1, \ldots, p_I \rangle, p' = \langle p'_1, \ldots, p'_I \rangle \in \mathcal{P}$, $T^0(p, p') = \prod_{i=1}^I T_i^0(p_i, p'_i|p_{-i})$.

We now incorporate mutations into this general evolutionary process. We specify for each population $i$ a mutation rate $\varepsilon \mu_i > 0$, with $\mu_i$ a population-specific parameter that governs the between-population relative frequency of mutations, and $\varepsilon$ an across-population parameter governing the overall frequency of mutations. We then alter the above no-mutation evolutionary process as follows: From a population state $p_t$ in period $t$, a preliminary (pre-mutation) population state $p_{t+1}^{(0)}$ for period $t + 1$ is chosen according to the transition probabilities $T^0$, i.e., according to the no-mutation evolutionary process.

This preliminary population state is then subjected to random mutations of the following form: in each population $i$, each member has probability $\varepsilon \mu_i$ of discarding her strategy and randomly selecting another from the strategy space $S_i$, with each strategy (including the one she just discarded) equally likely.\(^4\) This mutation process is carried out independently across the members of a population, and similarly across populations, resulting in the final population state for period $t + 1$, $p_t^{t+1}$. Notice that the expected number of mutations per generation in a population is proportional to the size of that population.

The evolutionary process with mutations can be summarized by the following scheme:

\[ p_t \rightarrow \text{selection (stochastic)} \rightarrow \text{mutation (stochastic)} \rightarrow p_t^{t+1}. \]

Within each population $i$, this is a stochastic process governed by the transition probabilities by $T_i^\varepsilon(p_i, p_i'|p_{-i})$. These individual population processes aggregate to an overall Markov process over the state space $\mathcal{P}$, defined by the transition probabilities

\(^4\)We can easily allow for the possibility that not all mutations between strategies within a population are equally likely; this case is discussed in Section 5.
\[ T^\varepsilon(p, p') = \prod_{i=1}^{I} T^\varepsilon_i (p_i, p'_i | p_{-i}) \] (because the independence of the within-population processes is not compromised by the mutations process we have defined).

Since \( \mu_i > 0 \) for each population \( i \), there is positive probability that, from any given population state, any other state can be reached (it just requires the appropriate mutations). Consequently, the evolutionary process \( T^\varepsilon(p, p') \) with positive mutation rates \( \mu_i \) is an ergodic Markov chain. It thus has a unique stationary distribution, which it approaches in the long run.

In principle, this stationary distribution is analytically calculatable, but in reality, for many games of interest, the state space will usually be so large that this calculation is infeasible. In general, the size of the state space, i.e., all possible population states, is

\[ |P| = \prod_{i=1}^{I} \binom{N_i+|S_i|-1}{|S_i|-1} \]

In the case of just two populations, each of size 20 members, and each with 4 strategies available to its members, the size of the state space is approximately \( 3 \times 10^6 \): calculating the stationary distribution of a Markov process over this state space thus involves solving a system of about \( 3 \times 10^6 \) linear equations. This problem intensifies as the population sizes increase.

In the next section, we shall employ a theorem from Fudenberg and Imhof (2006) to show that, when the mutation rate is very low for each population \( \varepsilon \ll 1 \), the stationary distribution of the evolutionary process with mutations, which is generally not efficiently computable, is in fact close to that of a simpler Markov process on a much-reduced state space, the set of all pure states \( P^{\text{pure}} \). The key idea behind this theorem is that, as the mutation rate approaches zero, the dynamics of the evolutionary process simplify: Starting from a pure state, in which each population is monomorphic, we wait a long time for a mutant to appear in any one of the populations. When it does, it subsequently either goes extinct or fixes in that population, typically before another mutant appears in any population, thereby re-establishing a pure state. Hence, the system spends almost all of its time in pure states, and we need only consider what happens when a single mutant appears in a single population.

The stationary distribution of this embedded Markov chain over the pure states, owing to the much smaller state space, will often be calculatable. (This reduced state space in the above example, with two populations of size 20, each of which has 4 available strategies, is of size 16. Notice too that the size of the reduced state space \( P^{\text{pure}} \) does not increase with increasing population size, unlike that of the full state space \( P \).

Moreover, we shall demonstrate that the asymmetry of the underlying game will render selection frequency-independent in the rare-mutations regime. This will make calculation of the transition probabilities of this reduced Markov chain significantly simpler than for symmetric games, allowing us to achieve sharper results in characterizing the stationary distribution of the embedded Markov chain.
3 The stationary distribution when mutations are rare

Assumptions 1 and 2, which concern the within-population no-mutation evolutionary processes $T^0_i$, translate into the following two straightforward propositions, stated without proof, concerning the aggregate no-mutation process $T^0$:

**Proposition 1.** Under $T^0$, all pure population states $p \in \mathcal{P}^{\text{pure}}$ are absorbing.

**Proposition 2.** Under $T^0$, all population states $p \in \mathcal{P}\setminus\mathcal{P}^{\text{pure}}$ are transient.

Let us label pure population states by $s = (s_1, \ldots, s_I) \in \mathcal{P}^{\text{pure}}$: here, all members of population $i$ play strategy $s_i \in S_i$. Denote by $s/s'_i$ the population state where every population $j \neq i$ is monomorphic for the strategy $s_j$, and population $i$ is monomorphic for the strategy $s_i$ except for one individual, who plays $s'_i \neq s_i$. Let the set of all such states, i.e., those where all populations but $i$ are pure, and $i$ is pure but for one individual, be $\mathcal{P}^{\text{pure}}/i$.

**Proposition 3.** Fix $s \in \mathcal{P}^{\text{pure}}$, and consider the limit $\lim_{\varepsilon \to 0} T^\varepsilon(s, p)$ for states $p \in \mathcal{P}\setminus\{s\}$. This limit exists for all states $p \in \mathcal{P}\setminus\{s\}$. However, $\lim_{\varepsilon \to 0} \frac{T^\varepsilon(s, p)}{\varepsilon} > 0$ if, and only if, $p \in \mathcal{P}^{\text{pure}}/i$ for some $i$. Otherwise, $\lim_{\varepsilon \to 0} \frac{T^\varepsilon(s, p)}{\varepsilon} = 0$.

To prove this, note that $T^\varepsilon(s, p)$ is a polynomial in $\varepsilon$ for all $p$. For $T^\varepsilon(s, s/s'_i)$, this polynomial has leading term $\frac{N_i \varepsilon \mu_i}{|S_i|}$, and so $\lim_{\varepsilon \to 0} \frac{T^\varepsilon(s, s/s'_i)}{\varepsilon} = \frac{N_i \mu_i}{|S_i|} > 0$. On the other hand, if the states $s$ and $p$ differ by the strategy played by more than one individual, then a one-step transition from the former to the latter requires more than one mutation to occur, and so $T^\varepsilon(s, p)$ has leading term of order $\varepsilon^k$, $k \geq 2$. Thus, for such states $p$, $\lim_{\varepsilon \to 0} \frac{T^\varepsilon(s, p)}{\varepsilon} = 0$.

Proposition 3 states that mutations from a pure state to a state where only one individual in one of the populations deviates from the pure state are, for very small $\varepsilon$ and thus very small mutation rates, at least an order of magnitude more likely than other transitions from the pure state (and, owing to the pure states being absorbing under the no-mutation process, mutations are the only way to transition out of pure states).

Now suppose that, from a pure state $s$, the system transitions to the state $s/s'_i$. Since interior states in population $i$ are transient, Assumption 2 implies that, absent further mutations in population $i$, the state shall be absorbed either back into the pure state $s$ (in which case we say that the mutant strategy $s'_i$ has ‘gone extinct’) or into the pure state $(s'_i, s_{-i}) = (s_1, \ldots, s_{i-1}, s'_i, s_{i+1}, \ldots, s_I)$ (in which case we say that the mutant strategy $s'_i$ has ‘fixed’).

But when the mutation rates are very small, we should expect this extinction or fixation of strategy $s'_i$ to occur before another mutant appears in population $i$, and indeed before
a mutant subsequently appears in any other population. This latter fact, that no mutant is expected to appear in any of the other populations during the extinction/fixation event in population  

populations on whose strategies the two pure states differ. This establishes positive probability the system can reach any other pure state in a number of steps equal to the number of populations on whose strategies the two pure states differ.

Population state

Thus, selection is frequency-independent in this regime, a fact that will make the calculation of the various fixation probabilities significantly simpler.

To formalize this intuition, for states  

pure states.\(^5\) Construct a  

K \times K transition probability matrix  

as follows:

- If the pure strategies labelled  

and  

respectively (i.e., pure states that differ by the strategy played by just one population), then  

- If the pure strategies labelled  

and  

differ by the strategy played by more than one population, then  

- Having thus defined  

for all distinct pairs  

and  

define  

\(\Lambda_{mn} = 1 - \sum_{n \neq m} \Lambda_{mn}\).

\(\Lambda\) is the transition probability matrix for a homogeneous Markov chain over the state space  

\(P\)\(\text{pure}\).\(^6\) Moreover, this Markov chain is irreducible, since any pure state can be reached from any other with positive probability in a number of steps equal to the number of populations on whose strategies the two pure states differ.

Finally, the Markov chain defined by  

is clearly recurrent: from any pure state, with positive probability the system can reach any other pure state in a number of steps equal to the number of populations on whose strategies the two pure states differ. This establishes the final proposition that we require, that  

induces a unique stationary distribution on the state space of pure population states (Karlin and Taylor, 1975):

\(^5\) A particular enumeration that we have found useful is where, writing  

\(K_i = \prod_{i=1}^{i+1} |S_i|\), the pure state  

is enumerated by  

\(\sum_{i=1}^{I-1} K_i(m_i - 1) + m_i\). The population states in the pure state enumerated  

can then be recovered as follows:  

\(m_i - 1 = n \mod |S_i|, \) and  

\(m_i - 1 = \left\lfloor \frac{n}{|S_i|} \right\rfloor \mod |S_i|\) for each  

\(^6\) If there are some  

such that  

\(\Lambda_{mm} < 0\) in the above construction of  

one can rescale all mutation rates  

by an appropriately small factor to render all  

\(\Lambda_{mm} > 0\). Any such rescaling will result in the same stationary distribution over the state space induced  

\(\Lambda\).
Proposition 4. There is a unique vector \( \lambda = (\lambda_1, \ldots, \lambda_K) \) such that \( \lambda_j \geq 0 \) for all \( j \), \( \lambda_1 + \ldots + \lambda_K = 1 \), and \( \lambda \Lambda = \lambda \).

We are now in a position to state our main result. Propositions 1-4 ensure that \( T^0 \), \( T^\varepsilon \), and \( \Lambda \) satisfy Assumptions 6-9 in Fudenberg and Imhof (2006). Employing their Theorem 2, we arrive at the following theorem.

**Theorem 1.** For each \( \varepsilon \), denote by \( \lambda^\varepsilon \) the unique stationary distribution of the Markov process \( T^\varepsilon \). If \( n \) corresponds, in the enumeration of pure states, to the pure state \( s \), then

\[
\lim_{\varepsilon \to 0} \lambda^\varepsilon(s) = \lambda_n.
\]

That is, the stationary distributions of \( T^\varepsilon \) approach \( \lambda \) as \( \varepsilon \) approaches zero, i.e., as mutation rates become small.

### 4 The usefulness of the result

Our result is useful on two fronts. First, it extends to asymmetric games our ability to compute stationary distributions of finite-population evolutionary or imitation processes. We have argued that it is these games, and asymmetric multi-stage games in particular, for which stochastic finite-population dynamics are most relevant.

Second, in the analogous result for a single evolving population interacting according to a symmetric game (Fudenberg and Imhof, 2006), the fixation probabilities that are used to construct the transition matrix \( \Lambda \) of the embedded Markov chain involve frequency-dependent selection (since fitnesses depend on the population state of this single population, which necessarily changes in the course of fixation or extinction of a mutant strategy within it). Unfortunately, for most well-known evolutionary or imitation processes, closed-form solutions (exact or approximate) for fixation probabilities either do not exist or are intractable when selection is frequency-dependent (Nowak, 2006).\(^7\)

On the other hand, we have shown that when the evolutionary process plays out over multiple populations interacting according to an asymmetric game, then the fixation probabilities used to construct \( \Lambda \) derive from frequency-independent selection. Since frequency-independent selection has long been a standard assumption of the population genetics literature, this fact allows us to make use of the many results about fixation probabilities in that literature. This bridge between evolutionary game theory and classical population genetics may allow for analytical calculation of the rare-mutations stationary distribution, where this would be impossible or infeasible in the typical single-population symmetric game setup.

---

\(^7\)Frequency-dependent fixation probabilities of certain processes can be shown to approach tractable representations as the population size becomes very large—see, e.g., Fudenberg and Imhof (2008). However, as we shall discuss in Section 5, the cases to which the `rare mutations' approximation studied here best applies are specifically those where population sizes are not too large.
To illustrate these points, we consider three examples of interest: a ‘battle of the sexes’ game, a discrete Crawford-Sobel signalling game (Crawford and Sobel, 1982), and the ‘beer-quiche’ game of Cho and Kreps (1987).

**Example 1: Battle of the sexes**

The well-known ‘battle of the sexes’ game involves a man and a woman hoping to coordinate their weekend activities, which are either going to a ballet performance (the woman’s preference) or going to a rugby match (the man’s preference). Both the man and the woman prefer coordination on either equilibrium to not coordinating. The simple example we shall study is summarized by the payoff matrix:

\[
\begin{array}{cc}
\text{Woman} & B & R \\
\text{Man} & B & 1, 2 & 0, 0 \\
& R & 0, 0 & 2, 1
\end{array}
\]

To cast this into an evolutionary model, assume two separate populations of men and women, of size \(N_m\) and \(N_w\) respectively. Each period, each member of each population goes either to the ballet or to the rugby match, and receives his/her expected payoff from interacting with a random member of the other population. (This corresponds to members of each group preferring to be at an event attended by many members of the other group, though males would prefer this to be at the rugby, and females would prefer it to be at the ballet—not too unworldly a scenario!)

Expected payoffs \(E\pi\) within both populations translate to fitnesses via the linear transformation \(f_\theta(E\pi) = 1 + \eta_\theta E\pi\), where \(\theta\) is either \(m\) (‘man’) or \(w\) (‘woman’), and \(\eta_m\) and \(\eta_w\) are the strengths of selection in the men’s and women’s populations respectively. The evolutionary, or imitation, dynamics within each population are assumed to be a Moran process (Moran, 1958; Nowak, 2006), occurring with mutations in the manner set out in Section 2. The per-person mutation, or experimentation/error, rates in the men’s and women’s populations are \(\varepsilon \mu_m\) and \(\varepsilon \mu_w\) respectively.

In populating the rare-mutations Markov matrix \(\Lambda\), we need only consider transitions between pure states where either the male population’s strategy is different or the female population’s strategy is different, but not both. For example, consider the transition from the pure state where the men all go to the rugby and the women all to the ballet, \((R, B)\), to the pure state where everyone goes to the rugby, \((R, R)\). (The bold font indicates that these are population strategies.) In the former ‘incumbent’ female population, members all had fitness \(1 + \eta_w(0) = 1\). A mutant woman going instead to the rugby has fitness \(1 + \eta_w(1) = 1 + \eta_w\), and thus has (frequency-independent) selective advantage \(\eta_w\) over the ballet-going members of the population.

This frequency-independence allows us to make use of the well-known formula for fixation probability under a Moran process: If a single mutant has selective advantage \(s\)
over the other members of the (size $N$) population, then its fixation probability is

$$\rho(s) = \frac{1 - 1/(1 + s)}{1 - 1/(1 + s)^N}$$

for $s \neq 0$, and $\rho(0) = 1/N$. The corresponding formula for the case of frequency-dependent selection is significantly more complicated (Ewens, 2004; Traulsen et al., 2006).

The entry of $\Lambda$ corresponding to the transition $(R, B) \rightarrow (R, R)$ is therefore

$$\Lambda_{(R, B) \rightarrow (R, R)} = \hat{\mu}_w \rho_w(\eta_w) = \frac{N_w \mu_w}{2} \frac{1 - 1/(1 + \eta_w)}{1 - 1/(1 + \eta_w)^N}.$$

For the reverse transition $(R, R) \rightarrow (R, B)$, mutant women have fitness 1, while incumbents have fitness $1 + \eta_w$. Mutants are thus at relative selective disadvantage $\frac{1}{1+\eta_w} = -\eta_w/(1 + \eta_w)$, and the relevant entry of $\Lambda$ is

$$\Lambda_{(R, R) \rightarrow (R, B)} = \hat{\mu}_w \rho_w\left(\frac{-\eta_w}{1 + \eta_w}\right) = \frac{N_w \mu_w}{2} \frac{1 - (1 + \eta_w)}{1 - (1 + \eta_w)^N}.$$

The other entries of $\Lambda$ are calculated similarly. Enumerating the pure states $(B, B)$, $(B, R)$, $(R, B)$, and $(R, R)$ as 1, 2, 3, and 4 respectively,

$$\Lambda = \begin{pmatrix}
1 - \ldots & \hat{\mu}_w \rho_w\left(\frac{-2\eta_w}{1+2\eta_w}\right) & \hat{\mu}_m \rho_m\left(\frac{-\eta_m}{1+\eta_m}\right) & 0 \\
\hat{\mu}_w \rho_w(2\eta_w) & 1 - \ldots & 0 & \hat{\mu}_m \rho_m(2\eta_m) \\
\hat{\mu}_m \rho_m(\eta_m) & 0 & 1 - \ldots & \hat{\mu}_w \rho_w(\eta_w) \\
0 & \hat{\mu}_m \rho_m\left(\frac{-2\eta_m}{1+2\eta_m}\right) & \hat{\mu}_w \rho_w\left(\frac{-\eta_w}{1+\eta_w}\right) & 1 - \ldots
\end{pmatrix},$$

where the ellipses abbreviate that the rows must each sum to one.

We have a number of free parameters in this model; to wit: the sizes of, selection strengths in, and mutation rates in the two populations. As an example, suppose we set the sizes of, and selection strengths in, the two populations equal at $N$ and $\eta$ respectively. Making use of the fact that, for the Moran process, $\rho(s)/\rho(-s/[1 + s]) = (1 + s)^{N-1}$, we calculate the stationary distribution of the Markov chain defined by $\Lambda$:

$$\lambda = \left[1, \frac{1}{(1 + 2\eta)^{N-1}}, \frac{1}{(1 + \eta)^{N-1}}, 1\right]/\tilde{\lambda},$$

where $\tilde{\lambda}$ is a normalization constant. Notice that, in the low mutation limit, the mutation rates, though possibly different in the two populations, do not affect the long-term distribution of states.

The proportions of time the populations spend both at the rugby and both at the ballet are equal, and are clearly seen to be higher for larger values of the common selection strength $\eta$ and population size $N$. The intuition for this effect of $\eta$ is straightforward: A higher $\eta$ increases the fixation probabilities of positively selected mutants, and decreases
the fixation probabilities of negatively selected mutants. In this coordination game, the
former are always mutants leading towards the coordination equilibria \((B, B)\) and \((R, R)\),
while the latter are always mutants leading away from these coordination equilibria.

The effect of population size can most easily be seen from the ratio of transition
probabilities from a non-coordination equilibrium to a coordination equilibrium (positive
selection \(s\)) and vice-versa (negative selection \(-s/(1+s)\)): \(\rho(s)/\rho(-s/(1+s)) = (1+s)^{N-1}\).
This ratio increases with \(N\), and so, for each path into and out of a coordination equi-
librium, a higher \(N\) increases the transition probabilities into, relative to the symmetric
probabilities out.

**Example 2: Crawford-Sobel signalling**

The next game to which we apply our result is a discrete variant of a signalling game from
Crawford and Sobel (1982, Sec. 4). Suppose that there are three possible states of the
world, \(\theta \in \{0, 1, 2\}\), with each equally likely. A signaller observes the state of the world,
and sends a costless signal \(s \in \{a, b, c\}\) to a receiver, who observes only the signal, and
not the state of the world. Having observed the signal, the receiver makes a decision \(r\).
Payoffs to signaller and receiver are as follows:

\[
\begin{align*}
\pi_S(\theta, r) &= -(r - \theta - \gamma)^2, \\
\pi_R(\theta, r) &= -(r - \theta)^2,
\end{align*}
\]

where \(\gamma \geq 0\) is a parameter that characterizes the signaller and receiver’s misalignment of
interests (for every \(\theta\), the receiver’s optimal decision is \(\gamma\) lower than the signaller would
most want it to be). For simplicity, we restrict the receiver’s possible decisions \(r\) to the
set \(\{0, 0.5, 1, 1.5, 2\}\), which covers all possible optimal decisions the receiver could make
given some posterior over the state space, having observed a signal.

For all \(\gamma \geq 0\), Nash equilibria exist where the signaller sends the same signal no matter
the state of the world, and the receiver, observing that signal, makes decision \(r = 1\). We call
these equilibria ‘uninformative’, and label them ‘xxx’, since the same signal \(x \in \{a, b, c\}\)
is sent for each state of the world \(\{0, 1, 2\}\). Also, for all \(\gamma\), Nash equilibria exist where the
signaller sends the same signal for states \(\theta = 0\) and \(\theta = 2\), and a different signal for state
\(\theta = 1\): to all sent signals, the receiver responds with decision 1. Since no practical (i.e.,
decision-changing) information is transmitted by the signaller, we also call these equilibria,
labelled ‘xyx’, ‘uninformative’.

For sufficiently low \(\gamma\), there also exist ‘partially informative’ equilibria where, for two
adjacent states of the world (i.e., \(\{0,1\}\) or \(\{1,2\}\)), the signaller sends the same signal,
but for the other state of the world, sends a different signal. For such ‘xxy’ and ‘xyy’
equilibria, these threshold values for \(\gamma\) are 0.1 and 0.75 respectively.

Finally, for \(\gamma \leq 0.5\), there exist ‘informative’ equilibria, where the signaller sends a
different signal for each state (‘xyz’), and the receiver makes a decision equal to the state
that the signal is sent from.
A full characterization of the equilibria of this game, including receiver’s responses to unsent signals required to sustain each equilibrium, is included in an appendix.

Crawford and Sobel (1982) argue, somewhat informally, that for a given value of $\gamma$, the most reasonable equilibria are the most informative ones possible for that $\gamma$. This, they claim, is because these equilibria are Pareto-superior to less informative equilibria, and are salient—or ‘focal’ in Schelling’s (1960) language—in that they are the most informative equilibria (the other salient equilibria are the least informative ones, but these are ruled out on the former grounds of being Pareto-inferior to the most informative equilibria).

The methodology developed in the present paper allows us to test this equilibrium selection prediction more formally, in the context of learning by agents. Notice that none of the equilibria that are not perfectly informative is strict, so that a deterministic infinite-population approach would be of little use here, particularly for higher values of the misalignment parameter $\gamma$ (for which the informative equilibria do not exist). Instead, our finite-population approach is better-suited to this game.

We assume two populations, one of signallers and one of receivers. The size of each population is $N$. Each signaller is equipped with a response to each possible state of the world, and each receiver with a response to each possible signal. States of the world are drawn independently for each individual interaction (i.e., there is no aggregate state of the world), and fitnesses are calculated according to expected payoffs.

Evolution within each population is assumed to be a Wright-Fisher process (Fisher, 1930; Wright, 1931), which has been used as a model for both biological evolution as well as imitation learning. We assume that expected payoffs translate to fitnesses exponentially, $f(E\pi) = \exp(\eta E\pi)$, with selection strength $\eta$ in both populations, and equivalent per-person mutation rates $\mu$.

In constructing $\Lambda$, frequency-independent selection allows us to make use of the well-known ‘diffusion approximation’ formula for the fixation probability, under the Wright-Fisher process, of a single mutant at selective advantage $s$ in a population of size $N$ (Kimura, 1962):

$$\rho(s) = \frac{1 - \exp(-s)}{1 - \exp(-Ns)}$$

for $s \neq 0$, and $\rho(0) = 1/N$. Again, the case of frequency-dependent selection is significantly more complicated (Lessard, 2005; Imhof and Nowak, 2006; Altrock et al., 2010).

We use these fixation probabilities to populate $\Lambda$ according to the method set out in Section 3, and calculate its stationary distribution. Fig. 1 plots, for the case $N = 100$ and $\eta = 1$, and for various values of the misalignment parameter $\gamma$, the relative frequencies of equilibria of different information levels in this stationary distribution.\(^8\)

It can be seen from Fig. 1 that the results of the learning/evolutionary dynamics in this
Figure 1: Frequencies of the signalling profiles of different levels of information transmission in the long-run dynamics of the Crawford-Sobel game, plotted for various values of the misalignment parameter $\gamma$. Both signaller and receiver populations are of size $N = 100$; fitness is exponential in expected payoffs, with equal selection strength $\eta = 1$; mutation rates are equal in the two populations. The results are broadly consistent with Crawford and Sobel’s prediction that the most informative equilibria supportable by a given value of $\gamma$ are the most reasonable for that $\gamma$.

game broadly support Crawford and Sobel’s prediction that the most informative equilibria supportable by a given $\gamma$ are the most reasonable for that $\gamma$. For low levels of misalignment $\gamma < 0.4$, the informative equilibria dominate, and information transmission is almost always perfect in the long run. For intermediate levels of misalignment ($0.4 < \gamma < 1$), partially informative equilibria, especially those of the form $xyy$, are dominant. For high levels of misalignment ($\gamma > 1$), only uninformative equilibria can be supported, and indeed such equilibria dominate the long-run dynamics.

Note that the equilibria involving signalling of the forms $xxy$ and $yxy$ do not have analogs in the equilibria of game with continuous state, signal, and decision spaces (Crawford and Sobel, 1982): they are artefacts of the discrete structure of the game we have set up. It is reassuring, then, that they play little role in the long-run dynamics, having very little frequency in the stationary distribution for all values of $\gamma$.

**Example 3: The beer-quiche game**

Our final example is the beer-quiche game of Cho and Kreps (1987), used there to illustrate the equilibrium refinement method they advance, the ‘Intuitive Criterion’. The extensive form of the game is given in Fig. 2. Player 1 is either a wimp (type $t_w$) or surly (type $t_s$).
Figure 2: Extensive form setup of the beer-quiche game of Cho and Kreps (1987).

\(t_s\), with probabilities 0.1 and 0.9 respectively. Player 1 knows his type; player 2 does not. Player 1 either has beer or quiche for breakfast, observed by player 2, who then chooses whether to fight player 1 or not. The payoffs are such that player 2 should choose to fight player 1 if the posterior probability he holds that player 1 is a wimp is greater than 0.5. For any action by player 2, player 1 prefers beer for breakfast if he is surly, but quiche if he is a wimp. Regardless of player 1’s type, he would prefer to avoid fighting.

The game has two Bayesian Nash equilibria, both of the ‘pooling’ kind: one in which player 1 eats quiche no matter his type, and one in which player 1 drinks beer no matter his type. In both cases, player 2 chooses not to fight in response to the observed behaviour of player 1, but would fight in response to the unobserved behaviour. Both pooling equilibria are sustained by player 2’s ‘out-of-equilibrium’ belief that, if he were to observe player 1 having the opposite breakfast to that consumed in equilibrium, there would be a greater-than-half chance that player 1’s type was wimp.

Cho and Kreps’s Intuitive Criterion, however, rules out the always-quiche equilibrium, by the argument that the out-of-equilibrium beliefs that player 2 is required to hold do not survive the following forward-inductive reasoning: If player 2 sees a deviation from quiche-eating, he should reason that player 1 must be surly, since a wimpish player 1 is achieving the highest possible payoff he can in the always-quiche equilibrium, and thus could have no incentive to deviate, whatever his beliefs about player 2’s likely response to such a deviation. Player 2 should therefore reason that a deviation by player 1 to beer-drinking signals that player 1 is surly, and, having reasoned thus, should not fight player 1 if he ever does deviate to drinking beer. But player 1, reasoning that player 2 should reason this way, should, in the event of being surly, deviate from the equilibrium; player 2 will not fight in response, and the surly player 1 will earn payoff 3, higher than the payoff he was earning in the proposed quiche-pooling equilibrium.

The Intuitive Criterion has been criticized as being, in some cases, too rationality-heavy (Fudenberg and Tirole, 1991). Our methodology allows us to test whether its prediction
in the beer-quiche game holds up under a rationality-light learning process, where players need not even know the other player’s payoffs.

We assume two populations, one for each role. Evolution proceeds in each population as a Wright-Fisher process with mutations. The population of player $i$’s, ‘population $i$’, is of size $N_i$, with selection strength $\eta_i$, exponential fitness $f_i = \exp(\eta_i E\pi)$, and per-individual mutation rate $\mu_i$. Each member of population 1 has a strategy prescribing his breakfast choice (beer or quiche) if he turns out to be wimpish (with probability 0.1) and if he turns out to be surly (with probability 0.9). Each member of population 2 has a strategy prescribing his response (fight or don’t fight) to seeing a member of population 1 drink beer for breakfast, and to seeing a member of population 1 eat quiche. Each round, each member of each population receives his expected ex-ante (i.e., before types are chosen in population 1) payoff from interacting with a random member of the other population.

We label pure population states by the tuple $b(t_w), b(t_s); r(B), r(Q)$: respectively, breakfast had when wimpish, breakfast had when surly; response to beer-drinking, response to quiche-eating. For the former two, $B$ and $Q$ represent ‘beer’ and ‘quiche’, while, for the latter two, $F$ and $N$ represent ‘fight’ and ‘no fight’. Again, the bold font is used to indicate that these are population strategies.

The weights of the most popular states in the stationary distribution are displayed in Fig. 3, for the parameter settings $\eta_1 = \eta_2 = 0.2$, $\mu_1 = \mu_2$, and for various population sizes $N = N_1 = N_2$. For large population sizes ($N > 20$), the pooling equilibrium predicted by the Intuitive Criterion, $BB; NF$, is the modal state in the stationary distribution. For all population sizes, the other pooling equilibrium, ‘all-quiche’, has low weight in the stationary distribution; this supports its rejection by the Intuitive Criterion.

Apart from the fact that, of the two Bayesian Nash equilibria, the one predicted by the Intuitive Criterion is dominant, it is also of interest that non-equilibrium population states occur so frequently in the long run. These states are, in order of their weights in the stationary distribution, $QB; NF$, $QB; NN$, and $BB; NN$. Indeed, for small population sizes ($N < 20$), the state $QB; NF$ has highest weight in the stationary distribution.

The success of these non-equilibrium states is a result of neutral and nearly-neutral drift. Starting from the equilibrium state $BB; NF$, members of population 2 who instead play $NN$ achieve the same expected payoff (0.9) as those playing $NF$, and so can neutrally invade the population. If they fix, the pure population state $BB; NN$ is established. From this state, members of population 1 who play $QB$ are slightly favoured over the incumbents playing $BB$ (expected payoff 3 versus 2.9), and so can invade and fix, establishing the pure state $QB; NN$. From this state, members of population 2 who play $NF$ are slightly favoured (expected payoff 1 versus incumbent expected payoff 0.9). If they invade and fix, pure state $QB; NF$ is established. But from this pure state, members of population 1 who play $BB$ are slightly favoured (expected payoff 2.9 versus incumbent 2.8). If they invade and fix, the equilibrium pure state $BB; NF$ is re-established. Notice that, because the reverse directions involve only neutral and slightly disfavourable mutations, they also occur with non-negligible probability, and are therefore likely to influence the stationary
Figure 3: The frequencies of various population states in the long run dynamics of the beer-quiche game, plotted for various common population sizes $N = N_1 = N_2$. For reference, population state ‘QB; NF’ is that where members of population 1 eat quiche (Q) if wimpish and drink beer (B) if surly, while members of population 2 do not fight (N) if they see beer-drinking and do fight (F) if they see quiche-eating. The equilibrium predicted by the Intuitive Criterion, BB; NF, is modal for large (>20), but not for low (<20), population sizes. The equilibrium ruled out by the Intuitive Criterion, QQ; FN, is infrequent in the long-run dynamics for all population sizes.

distribution.

The intuition for the fact that increased population size here results in the system spending more time in the Nash equilibrium state is similar to that for the same observation in the battle of the sexes. When the population size is small, mutants that are weakly selected against still have non-negligible probability of fixing, and so transitions out of BB; NF to, say, QB; NF (mutant’s expected payoff only 0.1 less than incumbents’) play a role in the long-run dynamics. When the population size is very low, mutants that are weakly selected against have very little chance of fixing, and so these paths out of equilibrium are shut down, leaving only neutral paths such as BB; NF $\rightarrow$ BB; NN. Increasing selection strength $\eta$ would have the same effect.

5 Discussion

Our model involves a number of assumptions and simplifications, three major ones of which we discuss below: (i) the assumption of multiple populations, one for each ‘role’ in the asymmetric game, (ii) that there are real situations of evolution and learning where mutations are sufficiently rare for the resulting evolutionary dynamics to resemble those
we have derived for the limiting case, and (iii) that mutation rates within populations are uniform. Thereafter, we discuss the relevance of the approach developed in this paper for mixed-strategy equilibria.

On the first, the assumption of multiple populations, an alternative approach would be to model evolution as occurring in a single population, wherein each agent has a strategy for every role. Expected payoffs to players would then be computed on the basis of random assignment of roles each period.

In most learning contexts, the multiple-population setup seems more natural: we think of roles as being assigned at the outset, with each agent subsequently learning how best to play her assigned role. A simple illustrative example is the battle of the sexes game studied in Section 4, where the gender of each agent is fixed for the duration of his/her learning period.

The multiple-population setup is also better suited to modelling the genetical evolution of multiple interacting, though reproductively distinct, species. In the context of genetical evolution within a single species, however, the more natural model is a single population of organisms, in whose genomes strategies for different roles are encoded at different loci (corresponding to different decision nodes in the underlying game). Strategies are then collections of alleles, one for each locus, and are inherited intact (ignoring recombination). In the course of the propagation of a strategy, which locus is relevant will change from generation to generation, as different roles are taken on (carrier is male or female, carrier is the incumbent occupant of a territory or the trespasser, etc.), and so it is most natural to treat selection as occurring on alleles at multiple loci in a single population.

When should we expect the evolutionary dynamics under this single-population multiple-loci (/multiple-role) model to resemble those under our multiple-population model (where each locus (/role) is treated as a separate 'population')? Here, the answer is simpler for deterministic infinite-population dynamics. If there is variation within the population for alleles (/strategies) at multiple loci (multiple loci exhibit ‘polymorphism’), then it can be shown that the multiple-population approach and the single-population approach yield equivalent dynamics under the deterministic replicator dynamics if a simple condition concerning allele frequencies holds (Cressman, 2003). This condition, known in the population genetics literature as linkage equilibrium, amounts to statistical independence of allele frequencies across loci, and is preserved through time under the replicator dynamics (Cressman, 2003).

In a finite population, polymorphism at multiple loci will be common if the mutation rate or the population size are sufficiently large. The stochastic nature of the evolutionary process in a finite population ensures that linkage equilibrium will not always hold, and so a ‘dynamical equivalence’ result such as that described above is not possible. Nonetheless, if mutations at different loci (i.e., at different decision nodes) occur independently, then in the regime of rare mutations studied in this paper, it will almost always be the case that at most one locus is polymorphic in the population. Thus linkage equilibrium will almost always hold, since linkage disequilibrium between two loci requires that both loci
be polymorphic. In the weak mutation limit, therefore, the dynamics are the same whether we model evolution as occurring in multiple populations of loci, or in a single multi-locus population.

On the second point, how rare do mutations have to be for the population dynamics to resemble those in our limiting case? A simple heuristic may be derived as follows: Assume all $I$ populations to be of size $N$, with a common individual mutation rate of $\mu$. Consider the case where, starting from a monomorphic state, a mutant appears in one of the population. Under most commonly studied population dynamics (e.g., Wright-Fisher, Moran), the typical time that it takes this mutant either to go extinct or fix in its population is of order $N$ or less, the longest being for the fixation of a neutral mutant (Kimura and Ohta, 1969; Ewens, 2004). Let us say that this time is $aN$. Then the probability that another mutant appears during the extinction/fixation of this mutant is of order $\mu N I \times a N$, and so, for this probability to be below some small threshold $\nu$, it is required that $\mu < \nu / (aN^2)$. If this holds, the dynamics should resemble those for the limiting case $\mu \to 0$.

Clearly, this case is most relevant either if mutations (or experimentation and errors) occur at a low per-period rate, or if the populations under study are small, or both. In learning dynamics, interpretation of this ‘rare mutations’ condition is difficult, since the rate of mutations is calibrated to the timescale over which strategy revisions are made. Thus, a ‘generation’ might in fact constitute a very short period of time, and we might expect experimentation or errors to be very infrequent on such a timescale. Interpretation of this condition is easier for genetical evolution, where the timescale is in generations, and the probabilities of mutations can be reasonably well measured. For example, the point mutation rate at a single nucleotide site in humans (though known to vary across the genome) is in the order of around $10^{-8}$ per generation (Roach et al., 2010; Lipson et al., 2015). If we set a threshold of $\nu = 0.05$ and $a = 1$, and consider evolution at two loci (‘roles’), then the above inequality holds for populations of up to about 1500 individuals.

On the third point, it may be objected that, in our model, mutation rates within populations are uniform: a mutation from any strategy to any other strategy is equally likely. While this assumption may be valid in certain genetic contexts, in a learning context we might expect certain errors, or examples of experimentation, to be less likely than others (Fudenberg and Levine, 1998). Also, in a genetical context, if we include in our concept of mutation the possibility of structural changes (e.g., rearrangements, translocations), or if we are interested in the evolutionary dynamics of a certain functional genotype relative to all other genotypes (grouped as one class), then asymmetric mutation rates would be natural (Nowak, 1992; McCandlish and Stoltzfus, 2014).

Our result can be generalized in a straightforward way to incorporate heterogeneity in mutation rates within populations. If we denote by $\varepsilon \mu_i(s_i, s_i')$ the probability that a member of preliminary period-$t$ population $i$ currently employing strategy $s_i$ will mutate to playing $s_i'$ in the finalized period-$t$ population, then the evolutionary process with mutations is a Markov chain $T^\varepsilon$. It is still the case that, for $s = \langle s_1, \ldots, s_I \rangle \in \mathcal{P}_{\text{pure}}$, ...
\[ \lim_{\varepsilon \to 0} T_{\varepsilon}(s,p) = 0 \text{ if } p \notin \mathcal{P}^{\text{pure}} \cup \mathcal{P}^{\text{pure}/i} \text{ for some } i. \] 
Now, however, for \( s/s' \in \mathcal{P}^{\text{pure}/i}, \)
\[ \lim_{\varepsilon \to 0} T_{\varepsilon}(s,s/s') = N_i \mu_i(s_i, s'_i). \] 
We can then construct the transition probability matrix \( \Lambda \) as we did before.

If it is always the case that \( \mu_i(s_i, s'_i) > 0 \), then the Markov chain defined by \( \Lambda \) is irreducible, and an analogous form of Theorem 1 goes through as before. If, however, we allow there to be some \( i, s_i, \) and \( s'_i \) such that \( \mu_i(s_i, s'_i) > 0 \), then \( \Lambda \) is no longer guaranteed to induce an irreducible Markov chain. It is required that it admit a unique stationary distribution \( \lambda \) for the analogous Theorem 2 to go through.

A final point concerns games with mixed-strategy equilibria. In evolutionary game theory, involving populations of agents, two kinds of ‘mixed strategy’ states must be distinguished (Grafen, 1979). The ‘population kind’ is where individuals within a population each play pure strategies, but different individuals play different strategies. In our setup, when mutations are rare, the system spends almost all of the long-run time in pure states (where individuals within each population all play the same strategy); mixed strategies of the ‘population kind’ are therefore essentially never observed. The underlying reason is that these polymorphic states are transient under the no-mutations process. For a different reason, these ‘population kind’ mixed states are also disallowed under the evolutionary stability concept of infinite-population deterministic dynamics: an evolutionarily stable mixed state would require the component strategies to have equal fitness, so that any of them could be involved in a ‘neutral invasion’ of the state, contradicting the possibility that the state is evolutionarily stable in the first place (Selten, 1980).

The second kind of mixed strategy state is the ‘individual kind’, and involves the individuals of a population all playing the same mixed strategy. Unlike the ‘population kind’, such states can be evolutionarily stable in infinite-population dynamics. They do, however, raise a problem for our finite-population approach. Allowing individuals to play any mixed strategy requires an infinite strategy space (the unit simplex in \( \mathbb{R}^{|S_i|} \), for population \( i \) with pure-strategy space \( S_i \)), and therefore an infinitely large state space. A workaround would be to approximate the infinite strategy space \( \mathbb{R}^{|S_i|} \) by a discrete lattice contained within it, although this is not entirely satisfactory, since mixed-strategy equilibria in asymmetric games are often very sensitive to perturbations from equilibrium proportions.

References

P. M. Altrock, C. S. Gokhale, and A. Traulsen. Stochastic slowdown in evolutionary processes. *Physical Review E*, 82(1):011925, 2010.

C. Cannings. The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Advances in Applied Probability*, pages 260–290, 1974.
I.-K. Cho and D. M. Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221, 1987.

V. P. Crawford and J. Sobel. Strategic information transmission. *Econometrica*, 50(6):1431–1451, 1982.

R. Cressman. *Evolutionary dynamics and extensive form games*. MIT Press, 2003.

J. F. Crow and M. Kimura. *An Introduction to Population Genetics Theory*. Harper & Row, 1970.

W. J. Ewens. *Mathematical Population Genetics. I. Theoretical Introduction*. Springer, 2004.

R. A. Fisher. *The Genetical Theory of Natural Selection*. Clarendon Press, 1930.

D. Foster and P. Young. Stochastic evolutionary game dynamics. *Theoretical Population Biology*, 38(2):219–232, 1990.

D. Fudenberg and L. A. Imhof. Imitation processes with small mutations. *Journal of Economic Theory*, 131(1):251–262, 2006.

D. Fudenberg and L. A. Imhof. Monotone imitation dynamics in large populations. *Journal of Economic Theory*, 140(1):229–245, 2008.

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

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

A. Grafen. The hawk-dove game played between relatives. *Animal Behaviour*, 27(3):905–907, 1979.

A. Grafen. Biological signals as handicaps. *Journal of Theoretical Biology*, 144(4):517–546, 1990.

L. A. Imhof and M. A. Nowak. Evolutionary game dynamics in a Wright-Fisher process. *Journal of Mathematical Biology*, 52(5):667–681, 2006.

M. Kandori, G. J. Mailath, and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993.

S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes*. Academic Press, 1975.

M. Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713–719, 1962.

M. Kimura and T. Ohta. The average number of generations until fixation of a mutant gene in a finite population. *Genetics*, 61(3):763, 1969.
F. E. Kydland and E. C. Prescott. Rules Rather than Discretion: The Inconsistency of Optimal Plans. *The Journal of Political Economy*, 85(3):473–492, 1977.

S. Lessard. Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theoretical Population Biology*, 68(1):19–27, 2005.

M. Lipson, P.-R. Loh, S. Sankararaman, N. Patterson, B. Berger, and D. Reich. Calibrating the human mutation rate via ancestral recombination density in diploid genomes. *bioRxiv*, 2015. doi: 10.1101/015560.

J. Maynard Smith. *Evolution and the Theory of Games*. Cambridge University press, 1982.

J. Maynard Smith and G. A. Parker. The logic of asymmetric contests. *Animal behaviour*, 24(1):159–175, 1976.

D. M. McCandlish and A. Stoltzfus. Modeling evolution using the probability of fixation: History and implications. *The Quarterly Review of Biology*, 89(3):225–252, 2014.

P. Milgrom and J. Roberts. Predation, reputation, and entry deterrence. *Journal of Economic Theory*, 27(2):280–312, 1982.

P. A. P. Moran. Random processes in genetics. *Proceedings of the Cambridge Philosophical Society*, 54(01):60–71, 1958.

M. A. Nowak. What is a quasispecies? *Trends in Ecology & Evolution*, 7(4):118–121, 1992.

M. A. Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.

J. Roach, G. Glusman, A. Smit, C. Huff, R. Hubley, P. Shannon, L. Rowen, K. Pant, N. Goodman, M. Bamshad, et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, 328(5978):636–639, 2010.

S. C. Salop. Strategic entry deterrence. *American Economic Review*, 69(2):335–38, 1979.

L. Samuelson. *Evolutionary games and equilibrium selection*. MIT Press, 1998.

L. Samuelson and J. Zhang. Evolutionary stability in asymmetric games. *Journal of Economic Theory*, 57(2):363–391, 1992.

T. C. Schelling. *The Strategy of Conflict*. Harvard University Press, 1960.

R. Selten. A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology*, 84(1):93–101, 1980.

A. M. Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, 1973.
A. Traulsen, J. M. Pacheco, and L. A. Imhof. Stochasticity and evolutionary stability. *Physical Review E*, 74(2):021905, 2006.

C. Veller, L. K. Hayward, A. Peysakhovich, M. Hoffman, and M. A. Nowak. Costly signalling emerges in evolutionary and learning dynamics. *Unpublished manuscript*, 2015.

S. Wright. Evolution in Mendelian populations. *Genetics*, 16(2):97–159, 1931.
Appendix table 1: Values of $\gamma$ for which the signalling equilibria of different information transmission levels exist. Sender strategies are of the form $(s(0), s(1), s(2))$, so that, for example, $xxy$ represents the strategies $abb$, $acc$, $baa$, etc. For each horizontal grey bar, the triplet inside is the receiver strategy $(r(x), r(y), r(z))$ that supports the Nash equilibrium the grey bar represents. The dark grey bar indicates strict Nash equilibrium.

| $\gamma$ | $0 < \gamma < 0.10$ | $\gamma = 0.10$ | $0.10 < \gamma < 0.5$ | $\gamma = 0.5$ | $0.5 < \gamma < 0.75$ | $\gamma = 0.75$ | $0.75 < \gamma < 1$ | $\gamma \geq 1$ |
|----------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-------------------|-------------------|
| xyz      | (0.0, 1.0, 2.0)  |                 |                   |                 |                   |                 |                   |                   |
| xxy      | (0.5, 2.0, 0.0)  |                 |                   |                 |                   |                 |                   |                   |
|          | (0.5, 2.0, 0.5)  |                 |                   |                 |                   |                 |                   |                   |
|          | (0.5, 2.0, 1.5)  |                 |                   |                 |                   |                 |                   |                   |
| xyr      | (1.0, 1.0, 1.0)  |                 |                   | (1.0, 1.0, 0.0) |                   | (1.0, 1.0, 0.5) |                   |                   |
| xyy      | (0.0, 1.5, 0.0)  |                 |                   |                 |                   |                 |                   |                   |
|          | (0.0, 1.5, 0.5)  |                 |                   |                 |                   |                 |                   |                   |
|          | (0.0, 1.5, 1.5)  |                 |                   |                 |                   |                 |                   |                   |
| xxx      | (1.0, 1.0, 1.0)  |                 |                   | (1.0, 1.0, 0.0) | (1.0, 0.0, 1.0)  | (1.0, 0.0, 0.0) | (1.0, 0.0, 0.5)  |                   |
|          | (1.0, 0.5, 0.5)  |                 |                   | (1.0, 0.5, 0.5) |                   |                 |                   |                   |
|          | (1.0, 0.5, 1.0)  |                 |                   | (1.0, 0.5, 1.0) |                   |                 |                   |                   |
|          | (1.0, 0.5, 1.0)  |                 |                   | (1.0, 0.5, 1.0) |                   |                 |                   |                   |
|          | (1.0, 0.5, 1.0)  |                 |                   | (1.0, 0.5, 1.0) |                   |                 |                   |                   |