Stochastic stability in binary choice coordination games

Mathias Staudigl*

European University Institute, Max Weber Programme, Via delle Fontanelle 10/20, I-50014 San Domenico di Fiesole (Italy)

Abstract

A recent literature in evolutionary game theory is devoted to the question of robust equilibrium selection under noisy best-response dynamics. In this paper we present a complete picture of equilibrium selection for asymmetric binary choice coordination games in the small noise limit. We achieve this by transforming the stochastic stability analysis into an optimal control problem, which can be solved analytically. This approach allows us to obtain precise and clean equilibrium selection results for all canonical noisy best-response dynamics which have been proposed so far in the literature, among which we find the best-response with mutations dynamics, the logit dynamics and the probit dynamics. Thereby we provide a complete answer to the equilibrium selection problem in general binary choice coordination games.

Keywords: Evolutionary game theory, Stochastic stability, Equilibrium selection, Bimatrix games
JEL: C72, C73

1. Introduction

The notion of Nash equilibrium is the most prominent solution concept in the theory of non-cooperative games. However, most games of interest in economic theory have multiple Nash equilibria, leaving open the question of which equilibrium we should regard as the most "relevant" one. Up to now there is no generally accepted theory of *equilibrium selection*, which is a large problem in many contexts where non-cooperative game theory is applied and empirically tested (see e.g. Bajari et al., 2010). To examine whether some outcomes in games are more likely than others, Foster and Young (1990), Kandori et al. (1993) and Young (1993) proposed dynamic models of play with persistent randomness. With the introduction of *noise* any outcome is obtainable in these models, but one can study the long-run likelihood that a certain outcome is obtained in a game when the noise term is tending to zero. An outcome is said to be *stochastically stable* if the long-run probability with which it is observed does not go to zero in the limit of vanishing noise. The central result of the aforementioned papers is that this intuitive concept not only rules out unstable mixed equilibra, but also allows us to select among *strict* Nash equilibria. These powerful results indicate that stochastic "evolutionary" models may serve as a building block for a theory of equilibrium selection.[1] Fascinated by this powerful machinery, many researchers have subsequently

*Phone:+39 055 4685 689, Fax:+39 055 4685 894
Email address: `mathias.staudigl@EUI.eu` (Mathias Staudigl)

[1]For alternative (deterministic) evolutionary dynamics which are able to select among strict Nash equilibria, see Matsui (1995), Hofbauer and Sorger (1999) and Oechssler and Riedel (2001).

extended and refined the predictions made by these models. However, the sober outcome of these studies is that we will not be able to build a robust theory of equilibrium selection with these methods. Two sources of non-robustness have mainly been identified by the literature. First, there is non-robustness to the specification of *noise* in the system (Bergin and Lipman, 1996). Second, the relative speed of adjustment of the players may influence equilibrium selection (Kandori et al., 1993, Alós-Ferrer and Netzer, 2010). Faced with these negative results we need to develop quantitative methods which directly examine the stochastic potentials of equilibria in order to foster our understanding of these models. The purpose of this paper is to present a promising approach to obtain such quantitative measures.

1.1. Outline of the paper

This paper presents a general analysis of the equilibrium selection problem in asymmetric binary choice coordination games under a general class of dynamic adjustment models, which have been recently introduced as *noisy best-response protocols* (Sandholm, 2010b). In a binary choice coordination game the two strict Nash equilibria are absorbing states under best-response processes without noise. Hence, once the players agree to coordinate on one convention, there is no way out of this convention. The introduction of noise, however, allows for small probability events in which the pattern of play deviates from the conventions, and therefore may trigger a sequence of player adjustments which may lead to the alternative convention. Let us label such paths *escape paths*. In the presence of small positive noise, it becomes a meaningful exercise to ask for the probability that the process follows an escape path. Finding the escape paths from both conventions with the highest probability allows us to compare the relative likelihood that the process wanders from one convention to the other. These paths can be interpreted as the *maximum likelihood* paths of escape. It is then intuitive to call one Nash equilibrium point *stochastically stable*, if the probability of its maximum likelihood escape path is lower than that identified for the competing equilibrium point.[2] Thus, the exercise we face is to find these maximum likelihood escape paths. It is well-known that these paths can be characterized by solving a dynamic program, i.e. a *shortest path problem* (see e.g. (Kandori et al., 1993, Kandori and Rob, 1995) and Young (1993; 1998)). In this problem formulation the states are identified as vertices in a weighted graph and the edge-weights are the *costs* of transiting from one state to another. Solving such a dynamic program is feasible via standard algorithmic methods. However, in order to obtain analytic results, this approach becomes rapidly infeasible. This is particularly true once different dynamics to the classical "mutation counting" (or *best-response with mutations*) dynamics are of interest. It seems that a different approach is needed to obtain a general characterization of maximum likelihood escape paths. In this paper we propose such an approach by deriving an optimal control problem whose solution will give us precisely a maximum likelihood escape path in the limit of large player sets. This method is indeed very powerful. For the first time it allows us to obtain a complete picture of the pattern of equilibrium selection in asymmetric binary choice coordination games under the three canonical perturbed best-response models: the best-response with mutations model of Kandori et al. (1993) and Young (1993), the logit choice model of Blume (1993), and the probit choice of Myatt and Wallace (2003). The classical results of Kandori et al. (1993) and Young (1993) tell us that the risk-dominant equilibrium is selected in (symmetric or asymmetric) coordination games, if the populations are of equal size.[3]

[2] This is also the idea behind the Radius-Coradius Theorem of Ellison (2000).

[3] In the "Darwinian dynamics" of Kandori et al. (1993) one has to add the assumption that the relative adjustment speed of the players is the same. In our model this would mean that the two player populations are of the same size.

We recover this result with our optimal control approach. In the logit choice model we show that the equilibrium which is *risk-dominant* in a suitably defined way, is uniquely stochastically stable. The probit choice model shows a more intricate selection pattern, and we observe that, whenever the coordination game has an exact potential, we obtain the selection of risk-dominant equilibria. In more general games this no longer holds. The evolutionary equilibrium selection of non-risk dominant equilibria in asymmetric games should be contrasted with results obtained by Blume (2003) and Sandholm (2010b). These authors have shown that for symmetric coordination games with linear incentives, e.g. population games with random matching of players, risk-dominant play is stochastically stable for any noisy-best response process. We show that this is no longer true in *asymmetric* games.

1.2. The optimal control approach

At this stage the reader may wonder what the above outlined optimal control approach buys us in the problem of equilibrium selection. Since the optimal control approach, developed in this paper, arises naturally from the problem formulation of evolutionary equilibrium selection, we would like to give an informal overview on the nature of the problem.

Stochastic evolutionary models are mainly concerned with the *asymptotic properties* of the game dynamics. Technically this means that we are interested in the structure of the invariant distribution of the Markov chain, which is (essentially) uniquely defined as rare events of experimentation render the process ergodic. This invariant distribution captures the long-run behavior of the players in two complementary ways. First it describes the long run behavior of almost all sample paths of the game dynamics. Second it is the limit distribution of the process as the number of rounds played goes to infinity. Thus, all the information about long-run play is stored in the invariant distribution. The literature on stochastic evolution in games has developed different approaches of equilibrium selection which can be paraphrased as the small noise limit, the large population limit, and their logically possible double limits. In this paper we focus on the small noise limit. For a state-of-the-art discussion of stochastic stability analysis the reader is referred to Sandholm (2010a), where a more detailed discussion of stochastic stability techniques can be found.

The small noise limit emphasizes the role of small probability events of experimentation by the players in a large population environment. Kandori et al. (1993) and Young (1993), building on the work of Freidlin and Wentzell (1998), used trees defined on the finite set of population states, to characterize those states that are played with non-negligible frequency as the noise level vanishes. Kandori and Rob (1995) subsequently demonstrated that this problem can be reduced to a dynamic program, which they solved for symmetric games of pure coordination and symmetric supermodular games. These optimization problems are defined on a discrete state space, the set of finite strategy distributions, and can be interpreted as shortest path problems by identifying the state space as a weighted directed graph. A drawback of this approach is that, as the population of players gets large, the shortest path problems become readily intractable. To overcome this "curse of dimensionality" problem, we show that the solution of the dynamic program converges to the solution of a continuous optimal control problem, which is solvable by standard methods. This convergence property is an important result, as it assures us that we can "trust" the solution of the optimal control problem

For the truncated fictitious play process of Young (1993) the standard bounds on sample and memory size must, of course, apply. Different population size can be interpreted as different sample sizes; See Hehenkamp (2003) for a more detailed discussion.

if the population of players is sufficiently large.[4] In fact the optimal control problem in this paper turns out to be fairly easy to solve, and it allows us to answer the central question we ask in this paper:

> How is equilibrium selection in general binary choice coordination games affected by the players' choice functions?

1.3. Related literature

This paper makes a contribution to a recent literature which is devoted to identifying the conditions for "detail-free" equilibrium selection in stochastic evolutionary models. To the best of my knowledge, this paper is the first attempt to investigate this question in asymmetric games. For symmetric binary choice coordination games, Sandholm (2010b) provides a complete picture of equilibrium selection under noisy-best response dynamics. He does not only investigate equilibrium selection in the small noise limit, but also considers the behavior of the model under the small noise and the large population limit. Moreover, he proves that these limits commute. We focus here on asymmetric games with two distinct player populations and the small noise limit. We obtain stochastic stability results by taking first the small noise limit and then the large population limit.[5] In this limiting scenario we capture the behavior of large populations of players who are playing slightly perturbed best-responses, i.e. with a large probability a revising player chooses a best-response, but there is a small chance of random experimentation. For economic analysis this limit is therefore particular interesting, as it models the play of populations of players who are with high probability expected utility maximizers, though in an essentially myopic way. The literature on learning and evolution in games has proposed many ways to rationalize this way of modeling game dynamics in economics.[6] It is an open question how the order of limits affects equilibrium selection in general, and we regard this as an important topic for future research.

The rest of the paper is organized as follows: In Section 2 we present the set of binary choice coordination games for which we can give complete equilibrium selection results. Section 3 presents the stochastic evolutionary model and defines noisy best-response functions. Section 4 is devoted to the problem of equilibrium selection. We report our results for the case of small player sets and large population sets, separately. Section 5 concludes.

2. Binary choice coordination games with large player sets

We consider population games in the spirit of Sandholm (2010a), specialized to two player populations and linear payoff functions. There are two populations of players $p \in \mathcal{P} = \{1, 2\}$. In each player population there are $N^p = Nm^p$ identical agents who are randomly matched with agents of the opponent population $3 - p$ to play this game. The numbers $N, N^p,\ p \in \mathcal{P}$ are integers, and $M = m^1 + m^2$ is the total mass of the society. We will call N, somewhat loosely, the *population size*. Let $\mathsf{X} := \mathsf{X}^1 \times \mathsf{X}^2 = [0, m^1] \times [0, m^2]$. Agents are assumed to use only

[4] Establishing precise error bounds is an interesting topic for future work.

[5] Hence, we prove equilibrium selection results in the *small noise double limit* in the sense of Sandholm (2010a). Inspiration for this approach came from the work by Kifer (1988; 1990) who studied random perturbations of discrete-time Markov chains satisfying a large-deviations principle.

[6] Fudenberg and Levine (1998) provides an intriguing discussion of learning models in economics.

pure actions, contained in the set $\mathcal{S}^p = \{s_1^p, s_2^p\}, p \in \mathcal{P}$. The aggregate distribution of behavior in population p is completely described by the frequency of players using action 1, denoted by $x^p \in \mathsf{X}^{p,N} := \{0, \frac{1}{N}, \ldots, \frac{N^p}{N}\}$. A population state is a pair $x = (x^1, x^2)$, living in the discrete grid $\mathsf{X}^N := \{x = (x^1, x^2) \in \mathsf{X} | Nx \in \mathbb{N}_0^2\} = \mathsf{X}^{1,N} \times \mathsf{X}^{2,N}$. The set X^N is an arbitrarily fine inner approximation of the rectangle X by choosing N sufficiently large. From our uniform random matching hypothesis it follows that the expected payoff of an agent in population $p \in \{1, 2\}$, when choosing action s_i^p, $p \in \{1, 2\}$, are linear, and given by the expressions

$$F_1^p(x^{3-p}) := a_p x^{3-p} + c_p(m^{3-p} - x^{3-p}),$$
$$F_2^p(x^{3-p}) := d_p x^{3-p} + b_p(m^{3-p} - x^{3-p}),$$

for numbers $(a_p, b_p, c_p, d_p) \in \mathbb{R}^4$. The essential information about the game is contained in the *incentive functions*

$$\begin{aligned} d^1(x^2) &:= F_2^1(x^2) - F_1^1(x^2) = \alpha^1(x_*^2 - x^2), \\ d^2(x^1) &:= F_2^2(x^1) - F_1^2(x^1) = \alpha^2(x_*^1 - x^1), \end{aligned} \tag{2.1}$$

where $\alpha^p := (a_p - d_p) + (b_p - c_p)$ and $x_*^p := m^p \frac{b_{3-p} - c_{3-p}}{\alpha^{3-p}}$ for $p \in \{1, 2, \}$. The dynamic models we consider in this paper only depend on the best-reply structure of the game. Therefore, we can normalize payoffs so that $\alpha^1 \equiv 1$ and $\alpha^2 \equiv \alpha$, and we can identify a population game with the vector field

$$d = (d^1, d^2) : \mathsf{X} \to \mathbb{R}^2, \ x = (x^1, x^2) \mapsto d(x).$$

Generically, the population game (d^1, d^2) has one or three Nash equilibria. Since the main concern in this paper is the question of evolutionary equilibrium selection, we will focus on the class of *coordination games*, which is characterized by the conditions

$$d((0, 0)) > 0 \text{ and } d((m^1, m^2)) < 0.$$

Following Harsanyi and Selten (1988), we call the *stability set* of strict Nash equilibrium A the rectangle $\mathsf{X}_A := [x_*^1, m^1] \times [x_*^2, m^2]$. Analogously, we call the stability set of equilibrium B the rectangle $\mathsf{X}_B := [0, x_*^1] \times [0, x_*^2]$. Figure 1 illustrates these concepts.

3. Stochastic evolution and stochastic stability

3.1. The stochastic evolutionary process

Our model of stochastic evolution builds on Hofbauer and Sandholm (2002; 2007) and Sandholm (2010b), but we adopt a discrete-time formulation as in Benaïm and Weibull (2003). At discrete points of time, contained in the set $\mathbb{T} := \{0, \frac{1}{N}, \frac{2}{N}, \ldots\}$, a single randomly chosen player receives a revision opportunity in which he considers changing his strategy. When a current i-player in population p receives a revision opportunity he switches to strategy $j \neq i$ with probability $\sigma^\eta(\pi)$, where $\pi \in \mathbb{R}$ represents the current payoff advantage of action j over i. The map $\sigma^\eta : \mathbb{R} \to (0, 1)$ is called a *choice function* and is parameterized by the noise level $\eta > 0$.[7] The composition $B_\eta^p := \sigma^\eta \circ d^p$ is a *noisy best-response function* if

$$\lim_{\eta \to 0} B_\eta^p(x^{-p}) = \begin{cases} 1 & \text{if } d^p(x^{-p}) < 0, \\ 0 & \text{if } d^p(x^{-p}) > 0. \end{cases} \tag{3.1}$$

[7] Our stochastic stability analysis can be extended to allow for heterogeneous choice functions by simply changing notation.

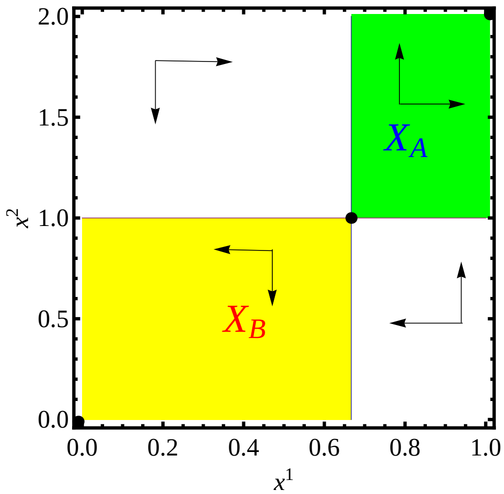

Figure 1: Illustration of a typical coordination game. Dots mark the position of Nash equilibria. Here we illustrate a continuum population game with two player populations with masses $m^1 = 1$ and $m^2 = 2$. Arrows indicate the direction of the vector field d. At the mixed equilibrium point x_* this vector field is trivial.

This dynamic process of play defines a finite state Markov chain $X^{N,\eta} = \{X^{N,\eta}_{k/N}\}_{k\in\mathbb{N}}$ on the lattice X^N with transition probabilities

$$\begin{aligned}\mathbb{P}\left(X^{N,\eta}_{(k+1)/N} = y | X^{N,\eta}_{k/N} = x\right) &= P^{N,\eta}(x,y)\\ &= \frac{1}{M}\begin{cases} (m^p - x^p)B^p_\eta(x^{-p}) & \text{if } y = x + \frac{1}{N}e_p,\ p \in \{1,2\},\\ x^p(1 - B^p_\eta(x^{-p})) & \text{if } y = x - \frac{1}{N}e_p,\ p \in \{1,2\},\end{cases}\end{aligned}$$

where e_1 and e_2 represent the canonical basis vectors of $\mathbb{R}^2$ and $t \in \mathbb{T}$. The set of feasible transitions of the Markov chain is denoted by $V = \{\pm e_1, \pm e_2\}$, and $v \in V$ represents a feasible direction of motion (i.e. a positively or negatively oriented unit vector). If $y \in \mathsf{X}^N$ but cannot be represented in the form $x + \frac{1}{N}v, v \in V$, then we set $P^{N,\eta}(x,y) = 0$.

3.2. Choice functions and their costs

The central concept of stochastic stability analysis is the notion of *costs* of a transition. Following Sandholm (2010b) we measure costs as the exponential rate of decay of the choice probability function in the small noise limit.

Hypothesis 1.

(1) There exists a function $w : \mathbb{R} \to \mathbb{R}_+$, to be called the waste function*, such that for all $\pi \in \mathbb{R}$*

$$-\lim_{\eta\to 0} \eta \log \sigma^\eta(\pi) = w(\pi) \tag{3.2}$$

with convergence uniform on compact intervals.

(2) The waste function $w : \mathbb{R} \to \mathbb{R}_+$ is said to be admissible *if*

(2.i) $w(\pi) > 0$ if $\pi > 0$,

(2.ii) $w(\pi) = 0$ if $\pi < 0$,

(2.iii) w is non-decreasing.

Note that we leave the waste function undefined at the point 0, i.e. we allow for $w(0) = 0$ and $w(0) > 0$. As observed by Sandholm (2010b) this creates more modeling freedom to allow for tie-breaking rules (see Example 1 below). Condition (1) says that the exponential rate of decay of a choice function is governed by the waste function w. Unpacking the expression (3.2) shows that for $\eta \to 0$

$$\sigma^{\eta}(\pi) = \exp\left(-\frac{1}{\eta}(w(\pi) + o(1))\right), \tag{3.3}$$

where $o(1)$ represents terms which are negligible as $\eta \to 0$. Any choice function whose waste function satisfies conditions (2.i) and (2.ii) generates a perturbed best response function $B^p_\eta = \sigma^\eta \circ d^p$. To see this note that Eq. (3.3) implies that

$$B^p_\eta(x^{-p}) = \exp\left(-\frac{1}{\eta}(w[d^p(x^{-p})] + o(1))\right)$$

as $\eta \to 0$. Thus, if strategy 1 is a best response for player p then $d^p(x^{-p}) < 0$ and by condition (2.ii) we must have $w(d^p(x^{-p})) = 0$. This implies that $B^p_\eta(x^{-p}) \to 0$ as $\eta \to 0$ in this case. But if strategy 2 is a best-response, then $d^p(x^{-p}) > 0$, and by condition (2.i) we see that $B^p_\eta(x^{-p}) \to 1, \eta \to 0$. Condition (2.$iii$) is a weak monotonicity assumption. We now present the canonical examples of noisy best-response functions, together with their waste functions.

Example 1 (Best-response with mutations)**.** *The most frequently applied model in stochastic evolutionary game theory assumes that players select with probability $1 - \epsilon$ a strategy that is strictly better compared to their current strategy. With complementary probability ϵ a player makes a random draw.*[8] *Parameterizing the mutation probability by $\epsilon = \exp(-1/\eta)$ gives the choice function*

$$\sigma^\eta(\pi) = \begin{cases} 1 - \exp(-1/\eta) & \text{if } \pi < 0 \\ \exp(-1/\eta) & \text{otherwise.} \end{cases}$$

For $\eta \to 0$ the waste function of this choice function is the mutation counting function

$$w(a) \equiv w_M(\pi) = \mathbb{1}_{[0,\infty)}(\pi). \tag{3.4}$$

In particular, we see that $w(0) = 1$.

◇

Example 2 (Logit choice)**.** *The logit choice function is defined as*

$$\sigma^\eta(\pi) = \frac{1}{1 + \exp(\pi/\eta)}.$$

For $\eta \to 0$ its waste function is given by

$$w(\pi) \equiv w_L(\pi) = \max\{0, \pi\}. \tag{3.5}$$

It is well known that this choice model can be derived from a random utility model with extreme-value distributed errors (see e.g. McFadden (1981) or Anderson et al. (1992)).

[8] This models a probabilistic choice behavior where agents are biased for their current strategy since a revising agent only makes a switch if the alternative promises a strictly higher (myopic) payoff. Alternative specifications of the choice model only changes the waste function at the point $\pi = 0$, which is possible by Hypothesis 1. This observation is due to Sandholm (2010b).

$\diamond$

Example 3 (Probit choice)**.** *Beside the Logit, another well known model from the discrete-choice literature is the Probit. While the logit choice model assumes extreme-value distributed errors, the probit choice function assumes normally distributed errors. Consider player* $p = 1, 2$ *and assume that*

$$\sigma^{\eta}(\pi) = \mathbb{P}(\epsilon_1 - \epsilon_2 > \pi)$$

where $\epsilon_j, j \in \{1, 2\}$*, are i.i.d.* $N(0, \eta/2)$ *distributed error terms. Denote by* $Z = \frac{\epsilon_1 - \epsilon_2}{\sqrt{\eta}} \sim N(0, 1)$*. Then it follows that*

$$\sigma^{\eta}(\pi) = \mathbb{P}\left(Z > \frac{\pi}{\sqrt{\eta}}\right) = 1 - \Phi\left(\frac{\pi}{\sqrt{\eta}}\right),$$

where Φ *is the cumulative distribution function of the standard normal distribution. Considering the limit* $\eta \to 0$*, it follows from the Laplace principle that the unlikelihood function is given by*

$$w(\pi) \equiv w_P(\pi) = \frac{\pi^2}{2}\mathbb{1}_{[0,\infty)}(\pi). \tag{3.6}$$

$\diamond$

4. Equilibrium selection

For finite population size N and positive noise level η the stochastic evolutionary process is a finite-state irreducible Markov chain. Standard results on finite Markov chains (see e.g. Stroock (2005)) tell us that for every parameter pair $(N, \eta) \in \mathbb{N} \times (0, \infty)$ the process $X^{N,\eta}$ possesses a unique invariant distribution, denoted by $\mu^{N,\eta}$. Since the pioneering work of Kandori et al. (1993) and Young (1993) it is common practice to call a state $x \in \mathsf{X}^N$ *stochastically stable* if it is contained in the support of the limiting invariant distribution $\mu^N := \lim_{\eta \to 0} \mu^{N,\eta}$. Since in this paper we work with more general choice functions, we are satisfied with a milder, large-deviations type, criterion of stochastic stability, which has been recently proposed by Sandholm (2010a;b).

Definition 1. *A state* $x \in \mathsf{X}^N$ *is called* stochastically stable in the small noise limit *if*

$$-\lim_{\eta \to 0} \frac{\eta}{N} \log \mu^{N,\eta}(x) = 0.$$

In this section we introduce some general methods with which stochastically stable states in the small noise limit and in the small noise large population double limit can be identified. A more formal account of the material presented here can be found in Appendix A.

4.1. Measuring stochastic stability

We define a *path* of length $K + 1$ as a sequence of states $\phi^N := (\phi^N_0, \phi^N_1, \ldots, \phi^N_K)$, where

(i) each $\phi^N_k \in \mathsf{X}^N$, and

(ii) $N(\phi^N_{k+1} - \phi^N_k) \in V$ for all $k \in \{0, \ldots, K - 1\}$.

Φ_x^N is the set of admissible paths starting at $x \in \mathsf{X}$. To measure the cost of a one-step transition along a path, we need to attach costs to every feasible direction of motion at a state $x \in \mathsf{X}^N$. Therefore we introduce the functions[9]

$$\begin{aligned} \kappa^p(x,v) &:= \operatorname{sgn}(v^p) w \left[\operatorname{sgn}(v^p) d^p(x^{-p})\right], \quad p \in \{1,2\},\ v \in V, \\ \kappa(x,v) &:= (\kappa^1(x,v), \kappa^2(x,v)) \end{aligned}$$

and define the *cost* of a transition from a point x in direction $v \in V$ as[10]

$$c(x,v) := \kappa^1(x,v)v^1 + \kappa^2(x,v)v^2 \equiv \langle \kappa(x,v), v \rangle . \tag{4.1}$$

The costs of a path ϕ^N are measured by the *action functional*

$$L_T(\phi^N) := \sum_{k=0}^{T-1} c[\phi_k^N, \phi_{k+1}^N - \phi_k^N],\ L_0 \equiv 0. \tag{4.2}$$

From the definition of waste functions it is clear that the process can leave the set $\mathsf{X}_{BA} = \mathsf{X} \setminus (\mathsf{X}_A \cup \mathsf{X}_B)$ at zero costs by letting the player for whom it is a best reply to do so adjust toward A or B (recall Figure 1). Every path that connects B with A (A with B) must incur the positive costs needed to wander through the set X_B (X_A) and no more. Hence, it is sufficient to define a path on X_B (X_A) to determine the costs of going from B to A (A to B). Moreover, it suffices to define an escape path only up to the point where it hits the outer boundary of the stability set, defined as

$$\begin{aligned} \partial\mathsf{X}_B &:= (\{x_*^1\} \times [0, x_*^2]) \cup ([0, x_*^1] \times \{x_*^2\}), \\ \partial\mathsf{X}_A &:= (\{x_*^1\} \times [x_*^2, m^2]) \cup ([x_*^1, m^1] \times \{x_*^2\}). \end{aligned}$$

However, due to finite population effects, these sets will in general have an empty intersection with the discrete grid X^N. To account for finite population effects we define for $p = 1, 2$ the smallest point on the grid X^N which is larger than or equal x_*^p as $\xi_N^p := \frac{1}{N}\lceil N x_*^p \rceil$. Similarly we call the largest points smaller than or equal x_*^p as $\zeta_N^p := \frac{1}{N}\lfloor N x_*^p \rfloor$. Then, we call $\partial\mathsf{X}_B^N = (\{\xi_N^1\} \times [0, \xi_N^2]) \cup ([0, \xi_N^1] \times \{\xi_N^2\})$ and $\partial\mathsf{X}_A^N = (\{\zeta_N^1\} \times [\zeta_N^2, m^2]) \cup ([\zeta_N^1, m^1] \times \{\zeta_N^2\})$.

Suppose we are given a mixed equilibrium $x_* = (x_*^1, x_*^2)$, a waste function w, and a population size parameter $N \geq N_0$. The *stochastic potentials* of the two strict Nash equilibria A and B are defined as

$$\gamma_A^N(w, x_*, \alpha) := \min\{L_K(\phi^N) | \phi^N \in \Phi_B^N,\ \phi_K^N \in \partial\mathsf{X}_B^N,\ K \in \mathbb{N}\}, \tag{4.3}$$
$$\gamma_B^N(w, x_*, \alpha) := \min\{L_K(\phi^N) | \phi^N \in \Phi_A^N,\ \phi_K^N \in \partial\mathsf{X}_A^N,\ K \in \mathbb{N}\}. \tag{4.4}$$

We define the *selection coefficient* as the difference in stochastic potentials, i.e.

$$S^N(w, x_*, \alpha) := \gamma_A^N(w, x_*, \alpha) - \gamma_B^N(w, x_*, \alpha). \tag{4.5}$$

A central result of stochastic stability analysis is that the strict Nash equilibrium with the smaller stochastic potential is uniquely stochastically stable (see e.g. Young, 1993; 1998). Thus, equilibrium

[9] For $\pi \neq 0$ we define $\operatorname{sgn}(\pi) := \frac{\pi}{|\pi|}$, and $\operatorname{sgn}(0) := 0$.

[10] In general we denote the inner product between two vectors x, y as $\langle x, y \rangle := x_1 y_1 + x_2 y_2$.

A is uniquely stochastically stable in the small noise limit at population size N if $S^N(w, x_*, \alpha) < 0$. If $S^N(w, x_*, \alpha) > 0$ then equilibrium B is uniquely stochastically stable in the small noise limit.
It is intuitively clear that in the population game (d^1, d^2) only the two strict Nash equilibria are candidates for the being stochastically stable in the small noise limit. Indeed, Young (1998, chapter 4) shows that this is true in his truncated fictitious play process. However, since we consider general noisy best-response dynamics, this still requires a proof.

Theorem 4.1. *Fix a population parameter $N \geq N_0$ and a waste function w. The set of stochastically stables states in the small noise limit in the population game (d^1, d^2) is contained in the set $\{(0,0), (m^1, m^2)\}$.*

Proof. See Appendix A. □

Hence, we have the following small noise limit characterization, whose proof can also be found in Appendix A.

Proposition 4.1. *For all $N \geq N_0$ we have*

$$\lim_{\eta \to 0} \left| \frac{\eta}{N} \log \frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)} - S^N(w, x_*, \alpha) \right| = 0. \tag{4.6}$$

This proposition endows us, in principle, with a general procedure for selecting between the two competing strict equilibrium points. In order to apply Proposition 4.1 one has to solve the dynamic program which underlies the definitions of the stochastic potentials $\gamma_A^N(w, x_*, \alpha)$ and $\gamma_B^N(w, x_*, \alpha)$. Then one obtains equilibrium selection results for every population size parameter N. However, the larger N the finer the grid $\mathsf{X}^N \subset \mathsf{X}$, and the more complex it becomes to solve the dynamic program lying behind the definition of the stochastic potentials. We circumvent this problem by following a common idea in evolutionary equilibrium selection, in which the action functionals are approximated by line integrals, and the dynamic programs are replaced by optimal control problems.[11] To do this in a meaningful way we have to invest some effort in making precise in which sense the equilibrium selection procedure for the finite state model is related to a suitably defined continuum model. This requires mastering some technical details which are collected in Appendix B. The prize for the effort is that we obtain precise and clean equilibrium selection results.
First we introduce the continuous counterpart of a path. In the continuum population model we call an absolutely continuous curve $\phi : [0, T] \to \mathsf{X}$ a *path*.[12] We endow $\mathbb{R}^2$ with norm $||x|| := |x^1| + |x^2|$. The sphere with respect to this norm is denoted by $U = \{x \in \mathbb{R}^2 | \; ||x|| = 1\}$ A path ϕ is admissible if for all t the tangent vector is contained in the set

$$\phi'(t) \in U(\phi(t)) := \{u \in U | (\exists \epsilon > 0) : \phi(t) + \epsilon u \in \mathsf{X}\}.$$

The continuous equivalent to the cost function (4.1) is then the 1-form

$$(\forall x \in \mathsf{X})(\forall u \in U(x)) : c(x, u) = \langle \kappa(x, u), u \rangle \, .$$

[11] Pioneers in this direction are Binmore et al. (1995), Binmore and Samuelson (1997), Maruta (2002), Blume (2003) and Sandholm (2010b).

[12] Recall that a function $\phi : [0, T] \to \mathsf{X}$ is absolutely continuous if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for any finite set of intervals $[t_0, t_1), [t_1, t_2), \ldots, [t_{n-1}, t_n)$ with $\sum_k (t_{k+1} - t_k) < \delta$ it follows that $\sum_k ||\phi(t_{k+1}) - \phi(t_k)|| < \epsilon$. In particular, an absolutely continuous curve is differentiable almost everywhere.

We define the line integral with respect to a piecewise continuously differentiable curve $\phi : [0, T] \to \mathsf{X}$ as

$$L(\phi) := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} c[\phi(t), \phi'(t)]dt,$$

where over each interval $[t_k, t_{k+1}) \subseteq [0, T]$ on which the curve ϕ is differentiable we evaluate the integral as

$$\int_{t_k}^{t_{k+1}} c[\phi(t), \phi'(t)]dt = \int_{t_k}^{t_{k+1}} \left(\kappa^1(\phi(t), \phi'(t))(\phi^1)'(t) + \kappa^2(\phi(t), \phi'(t))(\phi^2)'(t)\right) dt.$$

Now, let us focus on the problem of finding a least-cost path to exit the stability set X_B.[13] This can be formulated as the following variational problem:

$$\begin{aligned} & \min \textstyle\int_0^T c[x(t), u(t)]dt \\ \text{s.t.} \quad & x'(t) = u(t) \text{ a.a } t \in [0, T] \\ & u(t) \in U(x(t)) \text{ a.a } t \in [0, T] \\ & x(0) = B, \ x(T) \in \partial \mathsf{X}_B \\ & T \in [0, \infty) \end{aligned} \tag{4.7}$$

A solution is a pair $(\bar{\phi}(\cdot), T)$, consisting of an absolutely continuous curve $\bar{\phi} : [0, T] \to \mathsf{X}$, satisfying all the listed constraints. This curve is called a *least-cost path*, and the time point T is called the *exit time* of the least-cost path. Some comments are in order here. First, in the formulation of the optimization problem (4.7) we anticipate that an optimal control path for (4.7) will be non-decreasing, and non-increasing for the exit-problem from the set X_A. The reason for this is simple. Any non-monotonic path has segments with zero costs because it points into the direction of the equilibrium point we want to get away from. Without loss of generality we can cut out such segments. Second, we restrict the tangent vector of any feasible path to lie in the set $U(x)$. Also this is without loss of generality; Since any least cost path must be monotonic we can normalize the tangent vector to lie in the unit sphere without affecting optimality of the control.
Given this pre-information on the shape of a least cost exit path, we can write the running costs in problem (4.7), with a slight abuse of notation, as

$$\kappa(\bar{\phi}(t), \bar{\phi}'(t)) \equiv \kappa(\bar{\phi}(t)).$$

The value function of the optimal control problem measures the cost of an exit path, i.e.[14]

$$\gamma_A(w, x_*, \alpha) = \int_0^T c(\bar{\phi}(t), \bar{\phi}'(t))dt = \int_0^T \left\langle \kappa(\bar{\phi}(t)), \bar{\phi}'(t) \right\rangle dt.$$

As in the finite case, we declare the selection coefficient for the continuum problem to be the differences in the optimal value functions, i.e.

$$S(w, x_*, \alpha) := \gamma_A(w, x_*, \alpha) - \gamma_B(w, x_*, \alpha). \tag{4.8}$$

The following central result justifies our focus on large population equilibrium selection.

[13] The formulation of the least cost problem to exit X_A is formulated in a symmetric way.
[14] This is the optimal value function corresponding to problem (4.7).

Theorem 4.2.

$$\lim_{N\to\infty}\lim_{\eta\to 0}|\frac{\eta}{N}\log\frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)}-S(w,x_*,\alpha)|=0.$$

The proof of this theorem is rather long and technical so we present it in Appendix B. Theorem 4.2 tells us that in the small noise large population double limit the odds-ratio of the invariant distribution between the two strict Nash equilibria can be bounded as

$$\exp\left(-\frac{N}{\eta}S(w,x_*,\alpha)-\beta_1^{N,\eta}\right)\leq\frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)}\leq\exp\left(-\frac{N}{\eta}S(w,x_*,\alpha)+\beta_2^{N,\eta}\right)$$

for numbers $\beta_1^{N,\eta},\beta_2^{N,\eta}>0$ converging to 0 when we first take the small noise limit and then the large population limit. The practical implication of Theorem 4.2 is that it assures us that the selection coefficient defined in (4.8), obtained from the optimal control problem, really gives us an informative equilibrium selection criterion.

4.2. Main results

Having collected all the theoretical results, we are now ready to apply these tools to concrete equilibrium selection problems. We start with the relatively simple case in which there is one player in each player role. This is a standard normal form game in which two players repeatedly choose pure strategies according to the noisy-best response function B^p_η. Afterwards we turn to the analysis of the large population case.

4.2.1. Equilibrium selection for the case $N^p=1$.

If there is only one player in each player role the solution of the equilibrium selection problem is straightforward. The stochastic potential of the two Nash equilibria are[15]

$$\gamma_A(w,x_*,\alpha)=\min\{w(x_*^2),w(\alpha x_*^1)\}=w\left(\min\{\alpha x_*^1,x_*^2\}\right), \tag{4.9}$$

$$\gamma_B(w,x_*,\alpha)=\min\{w(\alpha(1-x_*^1)),w(1-x_*^2)\}=w\left(\min\{\alpha(1-x_*^1),1-x_*^2\}\right). \tag{4.10}$$

To see that these are indeed the costs of transitions, observe that the cost of a switch from action 1 to 2 is $w(-d^p(1))$, while the cost of a switch from action 2 to 1 is $w(d^p(0)),p\in\{1,2\}$. The relevant set of paths reduces to the set of sequences of the form $\phi=(\phi_0,\phi_1,\phi_2)$, with $\phi_0=(0,0)$ (all play strategy 2), $\phi_2=(1,1)$ (all play strategy 1), and $\phi_1\in\{(1,0),(0,1)\}$. Hence, to identify a least cost path only the first step of the path matters, and therefore the minimum function is used.

Observation 1. *The best-response with mutations dynamics of Example 1 does not allow us to make any prediction in this case. Therefore, in the rest of this section we assume that w is* monotonically increasing *on $\mathbb{R}_+$.*

In the coordination game with incentive functions (d^1,d^2), equilibrium A is $\frac{1}{2}$-dominant (Morris et al., 1995) if and only if $\max\{x_*^1,x_*^2\}<1/2$. It turns out that $\frac{1}{2}$-dominance is a sufficient condition for stochastic stability under any noisy best-response protocol.[16]

[15] These formulas hold, *mutatis mutandis*, also if the unlikelihood functions are heterogeneous, i.e. if the players use different noisy best-response protocols.

[16] Okada and Tercieux (2009) observe this for the logit choice.

Theorem 4.3. *Consider the coordination game (d^1, d^2) with $N^p = 1$. Then a $\frac{1}{2}$-dominant equilibrium is stochastically stable.*

Proof. We prove only the case where A is $\frac{1}{2}$-dominant. We have to distinguish the following cases. Suppose first that $\alpha \geq 1$. From $\max\{x_*^1, x_*^2\} < \frac{1}{2}$ we get $\min\{1 - x_*^1, 1 - x_*^2\} > \frac{1}{2}$. From the monotonicity of the waste function it follows that $\gamma_A(w, x_*, \alpha) = w\left(\min\{\alpha x_*^1, x_*^2\}\right) \leq w(x_*^2) < w(1/2)$, while $\gamma_B(w, x_*, \alpha) > w(1/2)$. Hence, $S(w, x_*) < 0$.
Now suppose that $\alpha < 1$. Now the expression $\min\{\alpha(1 - x_*^1), 1 - x_*^2\}$ can fall below $1/2$ if α is sufficiently small. Hence, we have to distinguish two cases. First let us assume that $x_*^2 > \alpha x_*^1$. Thus $\gamma_A(w, x_*, \alpha) = w(\alpha x_*^1)$. If $\gamma_B(w, x_*, \alpha) = w(\alpha(1 - x_*^1))$ then we see that $S(w, x_*, \alpha) = w(\alpha x_*^1) - w(\alpha(1 - x_*^1)) < 0$ as $x_*^1 < 1/2$. If $\gamma_B(w, x_*, \alpha) = w(1 - x_*^2) > \kappa(1/2)$ the claim follows from the fact that $w(\alpha x_*^1) < w(1/2)$. Second assume that $x_*^2 \leq \alpha x_*^1$. Then $\gamma_A(w, x_*, \alpha) = w(x_*^2)$ and $1 - x_*^2 \geq 1 - \alpha x_*^1 > \alpha(1 - x_*^2)$. Hence, $S(w, x_*, \alpha) = w(x_*^2) - w(\alpha(1 - x_*^1)) \leq w(\alpha x_*^1) - w(\alpha(1 - x_*^1))$ and the claim follows from the argument above. $\square$

Proposition 4.2. *Consider the coordination game (d^1, d^2) with $N^p = 1$ $p \in \mathcal{P}$. The implicit function $S(w, x_*, \alpha) = 0$ is solved by the piecewise linear curve*

$$x_*^2 = \sigma(x_*^1, \alpha) := \begin{cases} 1 - \alpha x_*^1 & \text{if } x_*^1 \in [0, \frac{1}{2\alpha}), \\ \frac{1}{2} & \text{if } x_*^1 \in [\frac{1}{2\alpha}, \frac{2\alpha - 1}{2\alpha}], \\ \alpha(1 - x_*^1) & \text{if } x_*^1 \in (\frac{2\alpha - 1}{2\alpha}, 1] \end{cases}$$

if $\alpha \geq 1$ and by

$$x_*^1 = \tilde{\sigma}(x_*^2, \alpha) := \begin{cases} 1 - \frac{x_*^2}{\alpha} & \text{if } x_*^2 \in [0, \frac{\alpha}{2}), \\ \frac{1}{2} & \text{if } x_*^2 \in [\frac{\alpha}{2}, \frac{2 - \alpha}{2}], \\ \frac{1 - x_*^2}{\alpha} & \text{if } x_*^2 \in (\frac{\alpha - 2}{\alpha}, 1]. \end{cases}$$

if $\alpha < 1$.

Proof. We only provide a verification of the formula for the case $\alpha \geq 1$. The case where $\alpha < 1$ is shown by almost identical arguments. For the proof we have to distinguish two cases.

(i) Assume that $\gamma_A(w, x_*, \alpha) = w(x_*^2)$. Then, by monotonicity, $x_*^2 \leq \alpha x_*^1$.

Case 1: If $\gamma_B(w, x_*, \alpha) = w(1 - x_*^2)$, then $x_*^2 \geq 1 - \alpha(1 - x_*^1)$. These two inequalities are consistent with $S(w, x_*) = 0$ if, and only if, $1/2 \in [1 - \alpha(1 - x_*^1), \alpha x_*^1]$, or if $x_*^1 \in [\frac{1}{2\alpha}, \frac{2\alpha - 1}{2\alpha}]$.

Case 2: If $\gamma_B(w, x_*, \alpha) = w\alpha(1 - x_*^1))$, then $x_*^2 \leq 1 - \alpha(1 - x_*^1)$. $S(w, x_*) = 0$ if, and only if $x_*^2 = \alpha(1 - x_*^1)$. This is consistent with the above inequality if, and only if $x_*^1 \geq \frac{2\alpha - 1}{2\alpha}$.

(ii) Assume that $\gamma_A(w, x_*, \alpha) = w(\alpha x_*^1) \Leftrightarrow \alpha x_*^1 < x_*^2$, by monotonicity of the cost function. Since $\alpha \geq 1$ it follows $\alpha(1 - x_*^1) \geq 1 - \alpha x_*^1$, and therefore $1 - x_*^2 < 1 - \alpha x_*^1 \Leftrightarrow \gamma_B(w, x_*, \alpha) = w(1 - x_*^2)$. Then $S(w, x_*, \alpha) = 0$ only if $x_*^2 = 1 - \alpha x_*^1$, which is consistent with the inequalities just established if, and only if, $x_*^1 \leq \frac{1}{2\alpha}$.

$\square$

A key quantity in stochastic stability analysis is the notion of *risk-dominance*. Let us view the state space X as a measure space endowed with Lebesgue product measure $\lambda := \lambda^1 \times \lambda^2$. An intuitive notion of risk-dominance requires that the stability set of one equilibrium covers a larger area in X than the stability set of the other equilibrium. Hence, we have the following definition:

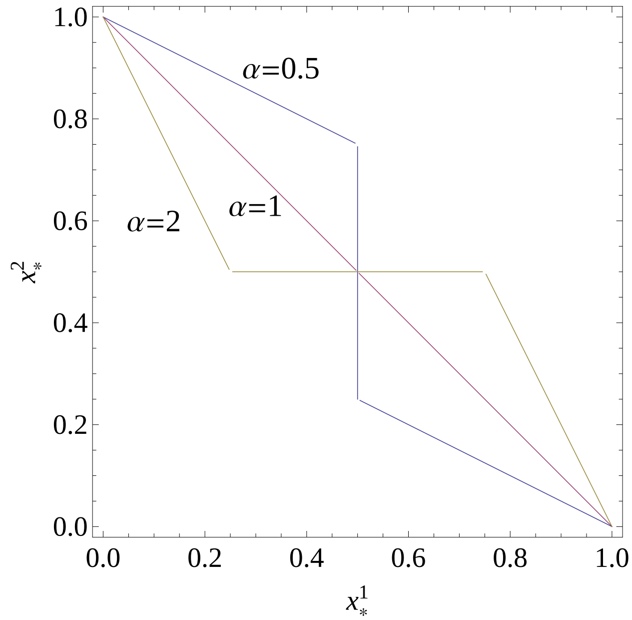

Figure 2: Plot of $S(w, x_*, \alpha) = 0$. The diagonal line arises for $\alpha = 1$. Below each of these lines equilibrium A is stochastically stable, while above each of these lines equilibrium B is stochastically stable.

Definition 2. *Strict Nash equilibrium A* risk-dominates *B if*

$$\lambda(\mathsf{X}_A) > \lambda(\mathsf{X}_B) \Leftrightarrow (m^1 - x^1_*)(m^2 - x^2_*) > x^1_* x^2_*.$$

Remark 1. *For $m^p = 1, p \in \mathcal{P}$, this is the standard definition of risk-dominance as given by Harsanyi and Selten (1988). For general population sizes our definition of risk-dominance is equivalent to the* stochastic dominance *concept of Sandholm (2010b), as our population games are characterized by linear incentives.*

Theorem 4.4. *If $\alpha = 1$ and $N^p = 1, p \in \mathcal{P}$, then a strict Nash equilibrium is stochastically stable under any noisy best-response protocol if, and only if, it is risk-dominant.*

Proof. If $\alpha = 1$ then $\sigma(x^1_*, 1) = 1 - x^1_*$. Hence $S(w, x_*, 1) > (<)0$ if and only if $x^1_* + x^2_* > (<)1$. $\square$

4.2.2. Equilibrium selection in large population environments

By construction of the revision dynamics we have $\dot{x}^1 + \dot{x}^2 = 1$. If we introduce the variable $u = \dot{x}^1$ we can write the variational problem (4.7) in the following, more structured, form:

$$\begin{aligned}
&\min_{u(\cdot),T} \int_0^T \left\{ w(x^2_* - x^2(t))u(t) + (1 - u(t))w(\alpha(x^1_* - x^1(t))) \right\} dt \\
\text{s.t.} \quad & \dot{x}^1(t) = u(t),\ \dot{x}^2(t) = 1 - u(t) \\
& u(t) \in [0, 1] \\
& x(0) = (x^1(0), x^2(0)) = (0, 0), \\
& x(T) = (x^1(T), x^2(T)) \in \partial \mathsf{X}_B, \\
& T \in [0, M]
\end{aligned} \tag{4.11}$$

Once we have solved this problem, finding the exit path from X_A is a relatively easy task. Suppose $\bar{\psi}$ is such an exit path from X_A. Let us perform the change of variables $\bar{\phi}^p(t) := m^p - \bar{\psi}^p(t)$ and denote by $y^p_* := m^p - x^p_*, p \in \mathcal{P}$. Then $(\bar{\phi}^p)'(t) = -(\bar{\psi}^p)'(t)$. Now, if $\bar{\phi}(\cdot)$ is an admissible path for the problem of finding a least cost path connecting equilibrium A with B it must be a *non-increasing* function, i.e. $(\bar{\phi}^1)'(t) = u(t) \in [-1, 0]$ and $(\bar{\phi}^2)'(t) = 1 - u(t)$. Hence, $(\bar{\psi}^p)'(t) = -u(t) \in [0, 1]$. Additionally, the set $\partial\mathsf{X}_A$ is transformed under this change of coordinates into the set $\partial\mathsf{Y}_A = \left(\{y^1_*\} \times [0, y^2_*]\right) \cup \left([0, y^1_*] \times \{y^2_*\}\right)$. Putting these insights together, we can formulate the optimal control problem to exit the set X_A with least cost as

$$
\begin{aligned}
&\min_{u(\cdot),T} \int_0^T \left\{ w(y^2_* - y^2(t))u(t) + (1 - u(t))w(\alpha(y^1_* - y^1(t))) \right\} dt \\
\text{s.t.} \quad & \dot{y}^1(t) = u(t), \ \dot{y}^2(t) = 1 - u(t) \\
& u(t) \in [0, 1] \\
& y(0) = (y^1(0), y^2(0)) = (0, 0), \\
& y(T) = (y^1(T), y^2(T)) \in \partial^+\mathsf{Y}_A, \\
& T \in [0, M]
\end{aligned}
\tag{4.12}
$$

This dynamic optimization problem is isomorphic to problem (4.11), and thus we only need to solve one of the two problems. In the rest of the paper we concentrate on the solution of (4.11).
To actually solve the optimal control problems we apply Pontryagin's Maximum principle. Let $\lambda := (\lambda_0, \lambda^1, \lambda^2)$ denote the vector of adjoint functions. The Hamiltonian can be written as

$$H(x(t), u(t), \lambda(t)) = I(x(t), \lambda(t)) + u(t)G(x(t), \lambda(t)),$$

where

$$
\begin{aligned}
I(x, \lambda) &:= \lambda^2 - \lambda_0 w(\alpha(x^1_* - x)), \\
G(x, \lambda) &: = \lambda_0 \left(w[\alpha(x^1_* - x^1)] - w(x^2_* - x^2)\right) + \lambda^1 - \lambda^2.
\end{aligned}
$$

As different degrees of convexity of the unlikelihood functions will generate different sets of results, we organize our presentation of the results accordingly. The interested reader may consult Appendix C for a full derivation of the solutions.

4.2.3. Strictly convex unlikelihood functions

Let us introduce the numbers

$$a(w; x_*, \alpha) := \alpha w'(\alpha x^1_*), \ b(w; x_*, \alpha) := w'(x^2_*). \tag{4.13}$$

These quantities are the marginal costs at $t = 0$ for any path in X which starts at the origin $(0, 0)$ and moves either in the direction of increasing x^1 or increasing x^2. If the players have unequal marginal costs, an optimal control path should exploit this and put full weight on the cost function of the player with lower marginal costs. This implies that $\bar{u}(t) \in \{0, 1\}$ must be optimal on some interval $[0, \tau)$. The point τ will depend on the data of the game and the unlikelihood function, and is defined as the time point at which a path hits the line on which the marginal costs of the transitions in the two feasible directions of motion are equalized. Once this is the case, we can calculate the rate of increase of the path in the direction of the x^1-axis as

$$g(x; x_*, \alpha) := \frac{w''(x^2_* - x^2)}{w''(x^2_* - x^2) + \alpha^2 w''[\alpha(x^1_* - x^1)]}. \tag{4.14}$$

Based on this intuition we have the following general result. Let $\mathbb{1}[s]$ be equal to 1 if statement s is true, and 0 otherwise.

Proposition 4.3. *Consider the optimal control problem (4.11) and suppose that the waste function w is strictly convex. Define*

$$\tau \equiv \tau(w; x_*, \alpha) := \begin{cases} x_*^2 - (w')^{-1}(a(w; x_*, \alpha)) & \text{if } a(w; x_*, \alpha) \leq b(w; x_*, \alpha) \\ x_*^1 - \frac{1}{\alpha}(w')^{-1}(b(w; x_*, \alpha)/\alpha) & \text{otherwise.} \end{cases} \tag{4.15}$$

The triple $(\bar{\phi}(\cdot), \bar{u}(\cdot), T)$ with

$$\begin{aligned} &\bar{u}(t) = \begin{cases} \mathbb{1}\left[a(w; x_*, w) > b(w, x_*, \alpha)\right] & \text{if } t \in [0, \tau), \\ g(\bar{\phi}(t); x_*, \alpha) & \text{if } t \in [\tau, T], \end{cases} \\ &T = x_*^1 + x_*^2, \\ &\bar{\phi}^1(t) = \int_0^t \bar{u}(s)\,\mathrm{d}\,s,\ \bar{\phi}^2(t) = t - \bar{\phi}^1(t),\ t \in [0, T], \end{aligned}$$

where $g(x; x_*, \alpha)$ is defined in equation (4.14), satisfies all the conditions provided by the Maximum principle with adjoint functions for $a(w; x_*, \alpha) \leq b(w; x_*, \alpha)$

$$\bar{\lambda}^1(t) = \begin{cases} a(w; x_*, \alpha)(\tau - t) + w(x_*^2 - \tau) & \text{if } t \in [0, \tau), \\ w(x_*^2 - \bar{\phi}^2(t)) & \text{if } t \in [\tau, T] \end{cases},\ \bar{\lambda}^2(t) = w[\alpha(x_*^1 - \bar{\phi}^1(t))]$$

or

$$\bar{\lambda}^1(t) = w(x_*^2 - \bar{\phi}^2(t)),\ \bar{\lambda}^2(t) = \begin{cases} b(w; x_*, \alpha)(\tau - t) + w(\alpha(x_*^1 - \tau)) & \text{if } t \in [0, \tau), \\ w(\alpha(x_*^1 - \bar{\phi}^1(t))) & \text{if } t \in [\tau, T] \end{cases}$$

for $a(w; x_*, \alpha) > b(w; x_*, \alpha)$.

Now, by the change of variables trick, we immediately get the reverse characterization of least cost paths to exit the stability set X_A. For sake of completeness we report our findings in the following Proposition.

Proposition 4.4. *Consider the optimal control problem (4.12) and suppose that the waste function w is strictly convex. Set $\tau \equiv \tau(w; m - x_*, \alpha)$ as defined in equation (4.15). The triple $(\bar{\psi}(\cdot), \bar{u}(\cdot), T)$ with*

$$\begin{aligned} &\bar{u}(t) = \begin{cases} \mathbb{1}\left[a(w; m - x_*, \alpha) > b(w; m - x_*, \alpha)\right] & \text{if } t \in [0, \tau), \\ g(m - \bar{\psi}(t); m - x_*, \alpha) & \text{if } t \in [\tau, T], \end{cases} \\ &T = M - x_*^1 - x_*^2, \\ &\bar{\psi}^1(t) = m^1 - \int_0^t \bar{u}(s)\,\mathrm{d}\,s,\ \bar{\psi}^2(t) = M - t - \bar{\psi}^1(t),\ t \in [0, T], \end{aligned}$$

where $g(m - x; m - x_*, w)$ is defined in equation (4.14), satisfies all the conditions provided by the Maximum principle with adjoint functions for $a(w; m - x_*, \alpha) \leq b(w; m - x_*, \alpha)$

$$\bar{\lambda}^1(t) = \begin{cases} a(w; m - x_*, \alpha)(\tau - t) + w(m^2 - x_*^2 - \tau) & \text{if } t \in [0, \tau), \\ w(\bar{\psi}^2(t) - x_*^2) & \text{if } t \in [\tau, T] \end{cases},\ \bar{\lambda}^2(t) = w[\alpha(\bar{\psi}^1(t) - x_*^1)]$$

or

$$\bar{\lambda}^1(t) = w(\bar{\psi}^2(t) - x_*^2),\ \bar{\lambda}^2(t) = \begin{cases} b(w; m - x_*, \alpha)(\tau - t) + w(\alpha(m^1 - x_*^1 - \tau)) & \text{if } t \in [0, \tau), \\ w[\alpha(\bar{\psi}^1(t) - x_*^1)] & \text{if } t \in [\tau, T] \end{cases}$$

for $a(w; m - x_*, \alpha) > b(w; m - x_*, \alpha)$.

The probit choice is the only canonical noisy best-response model with a strictly convex waste function. Thus, in order to get complete equilibrium selection results for the probit choice model we just have to make the correct substitutions into the formulas provided by Propositions 4.3 and 4.4. Recall from example 3 that the probit choice function has waste function $w_P(\pi) = \pi^2/2$ for $\pi \geq 0$. The stochastic potentials of the strict Nash equilibria can be computed as

$$\gamma_A(w_P, x_*, \alpha) = \begin{cases} \frac{(\alpha x_*^1)^2}{2}\left[x_*^2 - \frac{\alpha^2 x_*^1}{3}\right] & \text{if } x_*^2 \geq \alpha^2 x_*^1 \\ \frac{(x_*^2)^2}{2\alpha^2}\left[w^2 x_*^1 - \frac{x_*^2}{3}\right] & \text{otherwise.} \end{cases}$$

$$\gamma_B(w_P, x_*, \alpha) = \begin{cases} \frac{\alpha^2(m^1 - x_*^1)^2}{2}\left[(m^2 - x_*^2) - \frac{\alpha^2(m^1 - x_*^1)}{3}\right] & \text{if } m^2 - x_*^2 \geq \alpha^2(m^1 - x_*^1) \\ \frac{(m^2 - x_*^2)^2}{2\alpha^2}\left[\alpha^2(m^1 - x_*^1) - \frac{(m^2 - x_*^2)}{3}\right] & \text{otherwise.} \end{cases}$$

Figure 3 presents an example of an optimal control path for the Probit.

Proposition 4.5. *Consider a population game* (d^1, d^2) *with two equally sized populations of mass* $m^p = 1, p \in \mathcal{P}$. *If* $\alpha = 1$, *then an equilibrium is stochastically stable if and only if it is risk-dominant.*

Proof. The proof is a direct computation. The selection integral reduces in the case $\alpha = 1$ to

$$S(w_P, x_*, 1) = \left(\frac{1 - x_*^1 - x_*^2}{6}\right) P(x_*^1, x_*^2),$$

where $P(x, y) = -2 + x + x^2 + y - 4xy + y^2$ is a polynomial in (x, y) which attains a unique maximum at the point $(0, 1)$ at which it vanishes, and otherwise is negative. Hence $S(w_P, x_*, 1) < 0$ if and only if $x_*^1 + x_*^2 < 1$, and $S(w_P, x_*, 1) > 0$ if the reverse inequality holds. □

4.2.4. The best-response with mutations

In the best-response with mutations model the waste function is the mutation counting function $w_M(\pi) = \mathbb{1}_{\{\pi > 0\}}, \pi \in \mathbb{R}$. Hence, the criterion functional for problem (4.11) is

$$c(x, \dot{x}) = \mathbb{1}_{\{x^2 < x_*^2\}}\dot{x}^1 + \mathbb{1}_{\{x^1 < x_*^1\}}\dot{x}^2.$$

Let us consider the problem of how to direct a path in the most cost-efficient way away from the equilibrium point $B = (0, 0)$. By choosing an interior control, we select a path with initial cost equal to 1. Directing the path either along the vertical or horizontal axis of the state space gives the same costs. This is true at every point in the interior of X_B. However, choosing an interior control at some point inside this set increases the length of any path until it hits the target set $\partial\mathsf{X}_B$. Hence, no optimal control path can have an interior control at any point in $\operatorname{int}\mathsf{X}_B$. It follows that there are two candidates for a solution:

$$\phi_1(t) := (t, 0),\ t \in [0, T],\ T = x_*^1,$$

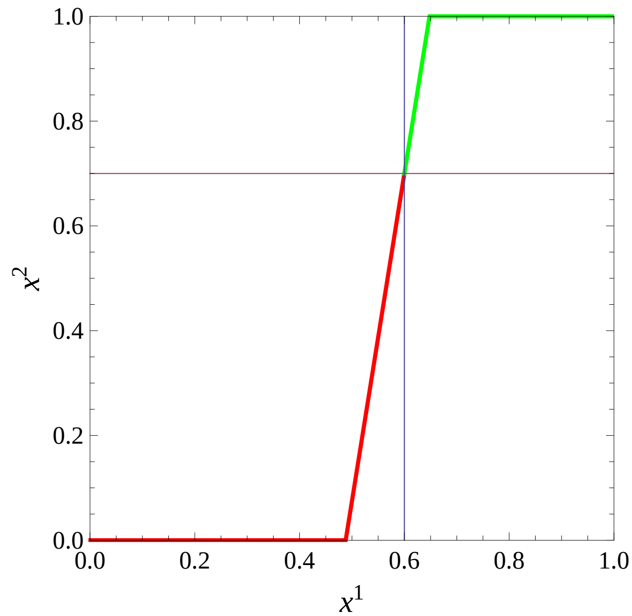

Figure 3: Optimal control paths for the probit choice function for a game between two populations of equal size with mixed Nash equilibrium $x_* = (0.6, 0.7)$ and $\alpha = 2.5$. In this game equilibrium $B \equiv (0, 0)$ is stochastically stable.

and

$$\phi_2(t) := (0, t),\ t \in [0, T],\ T = x_*^2.$$

The stochastic potential of equilibrium A is

$$\gamma_A(w_M, x_*, \alpha) = \min\{x_*^1, x_*^2\}.$$

In the same way we see that the stochastic potential of equilibrium B is

$$\gamma_B(w_M, x_*, \alpha) = \min\{m^1 - x_*^1, m^2 - x_*^2\}.$$

The selection coefficient reads as

$$S(w_M, x_*, \alpha) = \min\{x_*^1, x_*^2\} - \min\{m^1 - x_*^1, m^2 - x_*^2\}.$$

For $m^p = 1$, i.e. if both populations are of equal size, then we obtain the well-known relationship between risk-dominance and stochastic stability as first observed by Kandori et al. (1993) (see their section 9) and Young (1993), but proved with entirely different methods.[17] However, if relative population sizes differ, then this relationship breaks down.

Proposition 4.6. *In games with a risk-dominant equilibrium, the best-response with mutations dynamics selects the risk-dominant equilibrium if and only if $m^p = 1$.*

Proof. As soon as $m^p \neq 1$ for some $p \in \mathcal{P}$ the implicit function $S(w_M, x_*, w) = 0$ no longer defines a global function $x_*^2 = \sigma(x_*^1, w)$, but a sectionally smooth (linear) function in the rectangle X, such as in Section 4.2.1. $\square$

[17] To see this note that if $m^p = 1$ we have $x_*^1 < x_*^2 \Rightarrow 1 - x_*^1 > 1 - x_*^2$, and similarly $x_*^1 > x_*^2 \Rightarrow 1 - x_*^1 < 1 - x_*^2$. Hence, the selection coefficient takes the form $S(w_M, x_*, \alpha) = x_*^1 + x_*^2 - 1$.

4.2.5. The logit choice

The logit choice function has waste function $w_L(\pi) = \max\{\pi, 0\}$ (recall example 2). Consequently $w' = 1$ on $(0, \infty)$ and $(w')^{-1}$ is not a function (but a set-valued map). Hence, for the logit the solutions of Proposition 4.3-4.4 are not valid, and therefore this choice function has to be treated separately. Viewed from the origin of the rectangle X the waste functions w_L are distance functions for $(0,0)$ to the boundary points of the stability X_B, $(x^1_*, 0)$ and $(0, x^2_*)$. It is therefore intuitive that the optimal control path should be the same as the one obtained in the best-response with mutations case, i.e. we expect exit paths to be horizontal or vertical segments along the boundary of the rectangle X. To test this intuition in Proposition 4.7 we solve a series of auxiliary optimal control problems, where we choose as terminal state x^1_* or x^2_*. In both cases, the optimal control is indeed a straight path heading directly to these points. In Appendix C we then show that these are the only optimal control paths for the logit dynamics. As a result we obtain the following Proposition, which lists all cases exhaustively.

Proposition 4.7. *Consider the waste function of the logit choice function $w_L(\pi) = \max\{\pi, 0\}$.*

(a) The triple $(\bar{\phi}_1, \bar{u}_1, T_1)$, defined as

$$T_1 = x^1_*,\ \bar{\phi}_1(t) = (t, 0),\ \bar{u}_1(t) = 1,\ t \in [0, T_1],$$

solves the optimal control problem

$$\begin{aligned} &\min_{u,T} \int_0^T \left\{u(t)(x^2_* - x^2(t)) + (1 - u(t))\alpha(x^1_* - x^1(t))\right\} \mathrm{d}\, t \\ s.t.\quad & x(0) = (0,0),\ \dot{x}(t) = (u, 1-u), u \in [0,1] \\ & x^1(T) = x^1_*,\ x^2(T) \le x^2_* \end{aligned}$$

provided that $\alpha > 1$. The corresponding adjoint functions are

$$\bar{\lambda}^1(t) = x^2_*,\ \bar{\lambda}^2(t) = x^1_* - t,$$

and the stochastic potential is

$$\gamma_A(w_L, x_*, \alpha) = x^1_* x^2_*.$$

(b) The triple $(\bar{\phi}_2, \bar{u}_2, T_2)$, defined as

$$T = x^2_*,\ \bar{\phi}_2(t) = (0, t),\ \bar{u}_2(t) = 0,\ t \in [0, T_2],$$

solves the optimal control problem

$$\begin{aligned} &\min_{u,T} \int_0^T \left\{u(t)(x^2_* - x^2(t)) + (1 - u(t))\alpha(x^1_* - x^1(t))\right\} \mathrm{d}\, t \\ s.t.\quad & x(0) = (0,0),\ \dot{x}(t) = (u, 1-u), u \in [0,1] \\ & x^1(T) \le x^1_*,\ x^2(T) = x^2_* \end{aligned}$$

provided that $\alpha < 1$. The corresponding adjoint functions are

$$\bar{\lambda}^1(t) = -\alpha(t - x^2_*),\ \bar{\lambda}^2(t) = x^1_*,$$

and the stochastic potential is

$$\gamma_A(w_L, x_*, w) = \alpha x^1_* x^2_*.$$

(c) If $\alpha = 1$, then an optimal control is given by $\bar{u}_3(t) \equiv \bar{u} \in \{0,1\}$, with $T_3 = \max\{x_*^2, x_*^1\}$. The corresponding adjoint functions are

$$\lambda^1(t) = x_*^2 - x^2(t),\ \lambda^2(t) = x_*^1 - x^1(t).$$

The stochastic potential is given by

$$\gamma_A(w_L, x_*, \alpha) = x_*^1 x_*^2$$

Having solved the problem of connecting equilibrium B with the target set $\partial\mathsf{X}_B$, we can easily deduce the solution for the problem of finding the least cost path from equilibrium $A = (m^1, m^2)$ to the set $\partial\mathsf{X}_A$ by the change of variables trick. Hence, we obtain the following result:

Theorem 4.5. *Consider the coordination game (d^1, d^2) played by large populations with masses (m^1, m^2) where each player selects actions according to the log-linear choice rule. The selection integral is given by*

$$S(w_L, x_*, \alpha) = \min\{1, \alpha\} \left[x_*^1 x_*^2 - (m^1 - x_*^1)(m^2 - x_*^2)\right]. \tag{4.16}$$

This allows us to conclude:

Corollary 4.1. *A strict Nash equilibrium is uniquely stochastically stable under the log-linear choice rule if and only if it is risk-dominant according to definition 2.*

Proof of Theorem 4.5. From Proposition 4.7 we know that the stochastic potential of equilibrium $A \equiv (m^1, m^2)$ is given by

$$\gamma_A(w_L, x_*, \alpha) = \min\{1, \alpha\} x_*^1 x_*^2.$$

By performing the change of variables we conclude that the stochastic potential of equilibrium $B \equiv (0, 0)$ is given by

$$\gamma_B(w_L, x_*, \alpha) = \min\{1, \alpha\}(m^1 - x_*^1)(m^2 - x_*^2).$$

□

4.2.6. Application to symmetric binary choice coordination games

Our approach applies also to symmetric 2×2 games. To illustrate this, suppose that there is only a single population of players, where the fraction of 1-players is recorded by the one-dimensional state variable $x \in [0, 1]$.[18] In this case $\alpha = 1$ automatically holds, and we have an exact potential game. The selection integral takes the form

$$\begin{aligned} S(w, x_*, 1) &= \int_0^{x_*} w(x_* - t)dt - \int_{x_*}^1 w(t - x_*)dt \\ &= \int_0^1 \left[w(x_* - t) - w(t - x_*)\right] dt \\ &= \int_{x_*}^{1-x_*} \left[w(y) - w(-y)\right] dy \\ &= \int_{x_*}^{1-x_*} w(y)dy. \end{aligned}$$

[18]Taking the population mass to be one is without loss of generality here.

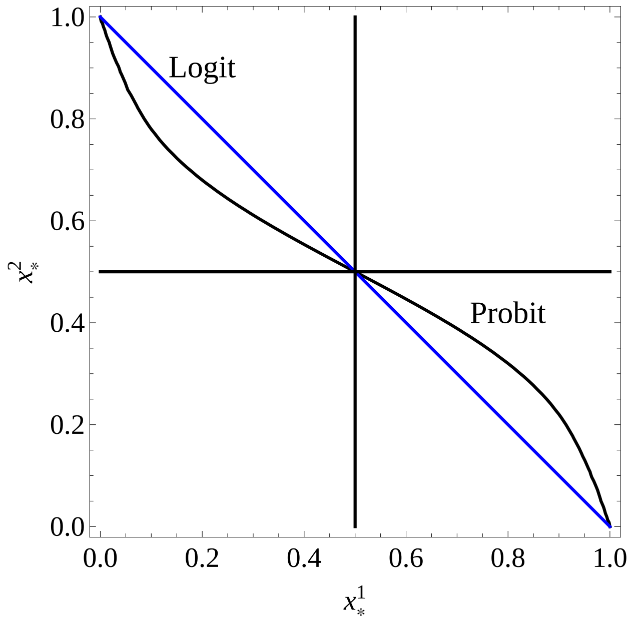

Figure 4: Plot of $S(w, x_*, \alpha) = 0$ for the logit and the probit for $\alpha = 3$. Below the curves equilibrium A is stochastically stable, while above the curve equilibrium B is stochastically stable.

From this expression it follows immediately that if $x_* > 1/2$ then $S(w, x_*, 1) < 0$ and equilibrium A is stochastically stable and risk-dominant. If $x_* < 1/2$ then equilibrium B is stochastically stable and risk-dominant. Finally if $x_* = 1/2$ no selection is possible according to either criteria. Symmetric binary choice coordination games have been more thoroughly investigated by Blume (2003) and Sandholm (2010b).

5. Conclusion

After all these calculations, it is time to discuss our findings. We have now obtained a complete picture of stochastic evolutionary equilibrium selection in asymmetric binary choice coordination games under arbitrary noisy best-response protocols. Our method of obtaining expressions for the stochastic potentials relied on a particular choice of taking limits. We calculate the "unlikelihood" that the dynamics cross the stability sets of the two Nash equilibria by defining a path cost function. In the large population limit we show that this path cost function can be approximated by a line integral. This allows us to formulate an optimal control problem to characterize the most likely path, paths we have called exit paths, among the unlikely ones. Evaluating the cost functional under this path gives us an expression for the stochastic potential of the two strict Nash equilibria. Since we are able to give a full characterization of exit paths, we regard our optimal control approach to stochastic evolutionary equilibrium selection as very valuable.
It has turned out that the stochastic potentials of Nash equilibria depend crucially on the degree of convexity of the waste functions. The role of convexity can be seen in the clearest way by looking at Figure 4. There we illustrate the regions of stochastic stability for the two canonical *random utility models*, the logit and the probit model. In this graph we plot the implicit function $S(w, x_*, 3) = 0$ for all positions of mixed equilibrium points x_* (we assume $m^p = 1, p \in \mathcal{P}$). In this case we know from our characterization in Proposition 4.7 that if players choose pure strategies according to

the logit choice function, then long-run play settles down to the risk-dominant equilibrium in the small noise limit. In stark contrast to this, the probit choice model displays a much more intricate pattern of equilibrium selection. In particular, we observe that for a large set of games, identified by the position of their mixed equilibrium $x_* = (x_*^1, x_*^2)$, the two models select different equilibria. In particular, what is selected by the probit protocol need not be a risk-dominant equilibrium in the sense of Harsanyi and Selten (1988). It is interesting to contrast this with recent results obtained by Sandholm (2010b). His Corollary 2 states that in symmetric binary choice coordination games obtained from random matching of players (i.e. when the incentive functions are linear) a Nash equilibrium is stochastically stable under any noisy best-response protocol if and only if it is risk-dominant. Proposition 4.5 extends this result to asymmetric binary choice games if the payoff parameters are chosen such that $\alpha = 1$, but not for $\alpha \neq 1$.

The current analysis has been simplified since we only deal with binary choice asymmetric coordination games with linear incentives. We leave it to future research to see how we can extend the optimal control approach to more general population games.

Acknowledgements

I am indebted to Josef Hofbauer and Gerhard Sorger for their astute comments and guidance. I also would like to thank David K. Levine and, in particular, William H. Sandholm for their very constructive comments on earlier versions of this work. The usual disclaimer applies. Financial support form the Vienna Science and Technology Fund (WWTF) under project fund MA 09-017, and the Max Weber Programme is gratefully acknowledged.

Appendix A Elements of stochastic stability analysis

For a fixed finite population size parameter $N \in \mathbb{N}$, we consider a family of perturbed Markov chains on the finite state space X^N, defined as

$$\mathcal{M}^N = \left(\Omega^N, (Y_k^\eta)_{k\in\mathbb{N}}, \mathcal{B}^N, \mathbb{P}^N\right)_{\eta\in\mathbb{R}_+}$$

where,

- $\Omega^N = (\mathsf{X}^N)^{\mathbb{N}}$ is the set of infinite sequences $x := (x_k)_{k\in\mathbb{N}}$ where each $x_k \in \mathsf{X}^N$,
- Y_k^η is the canonical projection mapping, i.e.

$$Y_k^\eta : \Omega^N \to \mathsf{X}^N,\ x = (x_k)_{k\in\mathbb{N}} \mapsto Y_k^\eta(x) = x_k$$

- $\mathcal{B}^N = \mathcal{B}(\Omega^N)$ the Borel σ-algebra on Ω^N.
- $\mathbb{P}^N$ is a probability measure on Ω^N associated with homogeneous Markov chains having transition probabilities $P^{N,\eta} : \mathsf{X}^N \times \mathsf{X}^N \to [0,1]$, defined by

$$\begin{aligned} P^{N,\eta}(x_{k-1}, x_k) &:= \mathbb{P}^N(Y_k^\eta = x_k | Y_0^\eta = x_0, Y_1^\eta = x_1, \dots, Y_{k-1}^\eta = x_{k-1}) \\ &= \mathbb{P}^N(Y_k^\eta = x_k | Y_{k-1}^\eta = x_{k-1}) \end{aligned}$$

  for any sequence $(x_0, \dots, x_{k-2}) \in (\mathsf{X}^N)^{k-1}$.

The basic assumptions on the transition probabilities of the Markov chain can be formulated in terms of a large deviations principle which forms the basis of this article.

Hypothesis 2. *(i) There exists a function $\rho^N : \mathsf{X}^N \times \mathsf{X}^N \to [0, \infty]$ such that for all $x, y \in \mathsf{X}^N$*

$$\lim_{\eta \to 0} \frac{\eta}{N} \log P^{N,\eta}(x, y) = -\rho^N(x, y). \tag{A.1}$$

(ii) The matrix

$$\left[\exp\left(-\frac{N}{\eta}\rho^N(x, y)\right)\right]_{(x,y)\in\mathsf{X}^N\times\mathsf{X}^N}$$

is irreducible.

It follows from these assumptions that for every fixed $N \geq N_0$ and $\eta > 0$ the Markov chain $\{Y_k^\eta\}_{k\in\mathbb{N}}$ has a unique invariant distribution $\mu^{N,\eta}$. The graph-theoretic approach of Freidlin and Wentzell (1998) allows one to represent this invariant distribution in an intuitive way. For every point $x \in \mathsf{X}^N$ call an x-graph $g \in G(x)$ a collection of arrows $(x' \to x'')$ such that (i) every point $x' \neq x$ is the origin of exactly one arrow, and (ii) the graph contains no cycle. Define

$$\theta_x^{N,\eta} := \sum_{g\in G(x)} \theta^{N,\eta}(g),\ \theta^{N,\eta}(g) := \prod_{(x'\to x'')\in g} P^{N,\eta}(x', x'').$$

Then one can show (see e.g. the Appendix of Young (1998), or Lemma 5.5, p.80, in Kifer (1988)), that the invariant distribution weights at $x \in \mathsf{X}^N$ as

$$\mu^{N,\eta}(x) = \frac{\theta_x^{N,\eta}}{\sum_{y\in\mathsf{X}^N} \theta_y^{N,\eta}}.$$

From this representation, and hypothesis 2, it is easy to see that

$$-\lim_{\eta\to 0} \frac{\eta}{N} \log \mu^{N,\eta}(x) = I^N(x) := W_x^N - \min_{y\in\mathsf{X}^N} W_y^N$$

for a bounded function $W^N : \mathsf{X}^N \to \mathbb{R}_+$ which is defined as

$$(\forall x \in \mathsf{X}^N) : \ W_x^N := \min\{W^N(g) | g \in G(x)\},$$
$$W^N(g) := \sum_{(x'\to x'')\in g} \rho^N(x', x'').$$

We immediately get the following general characterization of states which retain positive mass, at an exponential rate, in the limiting invariant distribution, proved by Kandori et al. (1993) and Young (1993).

Theorem A.1. *Any weak limit of the sequence of distributions $\{\mu^{N,\eta}\}_{\eta\geq 0}$ as $\eta \to 0$ is a probability distribution μ^N with support contained in the set $\{x \in \mathsf{X}^N | I^N(x) = 0\}$.*

Points in the set $(I^N)^{-1}(0)$ are population states appearing with highest probability at a logarithmic scale.

Definition 3. *The set* $\{x \in \mathsf{X}^N | I^N(x) = 0\}$ *is called the set of* stochastically stable states *in the small noise limit.*

We now apply these general results to our model of stochastic evolution. Therefore we set

$$X^{N,\eta}_{k/N} := Y^\eta_k,\ k \in \mathbb{N},$$

and identify the function ρ^N with

$$\rho^N(x,y) := \begin{cases} c(x, y-x) & \text{if } y-x \in \frac{1}{N}U^N(x), \\ +\infty & \text{otherwise.} \end{cases}$$

Recall the definition of the action functional (4.2), which associates to every path $\phi^N \in \Phi^N$ its total costs, i.e.

$$L_K(\phi^N) = \sum_{k=0}^{K-1} \rho^N(\phi^N_k, \phi^N_{k+1}) = \sum_{k=0}^{K-1} c(\phi^N_k, \phi^N_{k+1} - \phi^N_k),$$
$$L_0 \equiv 0.$$

For every two points $x, y \in \mathsf{X}^N$ let us introduce the (normalized) cost function

$$\frac{1}{N}C^N(x,y) = \inf\{L_K(\phi^N) | \phi^N \in \Phi^N_{x,y}, K \geq 1\}. \tag{A.2}$$

We collect some properties of the cost function $C^N(x,y)$, which are easy to verify.

Lemma A.1. *The cost function (A.2) has the following properties:*

(i) For every $x, y \in \mathsf{X}^N$ *we have that* $C^N(x,y) \in [0,\infty)$.

(ii) If $x \notin \mathsf{X}^N_A$ *then* $C^N(x,B) = 0$, *and if* $x \notin \mathsf{X}^N_B$ *then* $C^N(x,A) = 0$.

Hence, from any point outside the stability set of a strict equilibrium we can find a path which connects that point with the other strict equilibrium point at zero costs. Therefore, let us introduce the sets

$$\mathcal{D}^N_\sigma := \{x \in \mathsf{X}^N | C^N(x,\sigma) = 0\}, \quad \sigma \in \{A, B\}. \tag{A.3}$$

We can then prove the following fact, which also proves Theorem 4.1.

Lemma A.2. $\{x \in \mathsf{X}^N | I^N(x) = 0\} \subseteq \{A, B\}$.

Proof. If $x \notin \{A,B\}$ then, by Lemma A.1, $C^N(x,A) = 0$ or $C^N(x,B) = 0$. Suppose $C^N(x,A) = 0$. The case where $C^N(x,B) = 0$ is analogous and therefore omitted. Consider a least cost x-graph $g^*_x \in G(x)$ and let $\phi^N_{A,x}$ represent the unique path connecting $A = (m^1, m^2)$ with the point x on the least cost graph g^*_x. Delete this path and add a zero cost path $\phi^N_{x,A} \in \Phi^N_{x,A}$. This is possible by hypothesis. Hence,

$$W^N_A \leq W^N_x - C^N(A,x) + C^N(x,A) = W^N_x - C^N(A,x) < W^N_x.$$

Hence $x \in \mathcal{D}^N_A \Rightarrow W^N_A < W^N_x$. Similarly it follows that $x \in \mathcal{D}^N_B \Rightarrow W^N_B < W^N_x$. Hence, for all $x \notin \{A,B\}$ it follows that $I^N(x) > 0$. $\square$

This shows that, in order to single out predictions which are expected to be the most persistent ones, we are asked to compute the numbers W_A^N, W_B^N. These quantities are known in the game-theoretic literature as the stochastic potentials of the equilibrium points A and B. This shows that our analysis has to focus on the computation of the numbers W_A^N and W_B^N.[19] In the following Lemma we identify these functions as the well-known stochastic potentials, already introduced in (4.3) and (4.4).

Lemma A.3. $W_\sigma^N = \gamma_\sigma^N(w, x_*),\ \sigma \in \{A, B\}$.

Proof. We proof this Lemma only for the case $\sigma = A$. The case for the other equilibrium point is proved analogously, by only changing notation. We need to show that the costs accumulated on a least cost graph $g_A \in G(A)$ equals the costs accumulated on the least cost path $\phi_A^N \in \Phi_B^N$. Therefore, by Lemma A.1, we only need to consider states in X_B^N, as from all other states we find a null cost path which connects them to A. However, on X_B^N the graph g_A cannot have higher costs than the path ϕ_A^N, since otherwise we could form a path with lower costs, leading to a contradiction. ☐

Hence, equilibrium selection is determined entirely by the selection coefficient

$$S^N(w, x_*, \alpha) = \gamma_A^N(w, x_*, \alpha) - \gamma_B^N(w, x_*, \alpha).$$

Let us now have a closer look at the programs lying behind the value functions $\gamma_\sigma^N(w, x_*),\ \sigma \in \{A, B\}$. For sake of exposition let us focus on the the stochastic potential of equilibrium point A. It is by now clear that stochastic potentials are the value functions of the following dynamic program:

$$\begin{array}{rl} & \min \sum_{k=0}^{K-1} c(x_k, u_k) \\ \text{s.t.} & x_0 = B,\ x_K \notin \mathsf{X}_B \\ & x_{k+1} = x_k + \frac{1}{N} u_k \\ & u_k \in U^N(x_k)\ \forall k = 0, 1, \ldots, K-1 \\ & K \in \mathbb{N} \end{array} \tag{A.4}$$

where $U^N(x) := \{u \in V | x + \frac{1}{N} u \in \mathsf{X}^N\}$ is the set of feasible directions (the control variable) at state $x \in \mathsf{X}^N$. The dynamic program has a solution (it is a finite problem), which we denote as a pair $(\bar{\phi}_{BA}^N, K_{BA}^N)$. The (finite) sequence $\bar{\phi}_{BA}^N$ is a *least-cost path* for the finite-state Markov chain $X^{N,\eta}$.[20] The optimal value function of the dynamic program is exactly the action functional, evaluated under a least-cost path. Since this is exactly the stochastic potential of a strict Nash equilibrium, we have

$$\gamma_A^N(w, x_*) = L_{K_{BA}^N}(\bar{\phi}_{BA}^N),\ \gamma_B^N(w, x_*) = L_{K_{AB}^N}(\bar{\phi}_{AB}^N).$$

The following Lemma gives a preliminary characterization of least cost paths.

Lemma A.4. *Fix $N \geq N_0$ and consider the dynamic program (A.4). Then in order to wander from point A to B (B to A) only non-increasing (non-decreasing) paths can be least-cost.*

Proof. Focus on the problem to exit X_B. From Lemma A.1 we only have to consider paths with image on this set. But then any path which is not monotonic only makes the path longer, without adding something to the performance. Since the length of a path is part of the minimization problem, we can discard such paths. ☐

[19] In particular, Lemma A.2 says that the mixed equilibrium point x_*, whenever it lies on the grid X^N, cannot be a candidate for a stochastically stable state.

[20] The associated control sequence can be, of course, reconstructed by looking at the increments of the least-cost path.

Appendix B The large population limit

In this section we will show that solutions of a suitably defined variational problem are intimately related to solutions of the dynamic program (A.4). To introduce a continuum approximation we have to define our class of paths upon which we will then formulate the optimization problem. We endow $\mathbb{R}^2$ with the norm $||x|| := |x^1| + |x^2|$ for every $x \in \mathbb{R}^2$. The 1-sphere with respect to this norm is the set $U := \{x \in \mathbb{R}^2 | \, ||x|| = 1\}$. For every point $x \in \mathsf{X}$ we define the set of admissible directions of motion as the tangent space at x

$$U(x) := \{u \in U | x + \epsilon u \in \mathsf{X} \text{ for some } \epsilon > 0\}.$$

Note that for every $N \geq N_0$ and $x \in \mathsf{X}^N$ we have $U^N(x) \subseteq U(x)$. Now, we need to declare a continuous equivalent of the running cost index (4.1). For every tangent vector $u \in U(x)$ a natural extension of the running cost index of the finite problem is the mapping[21]

$$c(x, u) := \langle \kappa(x, u), u \rangle \tag{B.1}$$

The task of this section is to make it precise how and if the selection coefficients of the finite problem are related to the selection coefficient of the continuum problem. This is achieved by the following Theorem.

Theorem B.1.

$$\lim_{N \to \infty} |S^N(w, x_*, \alpha) - S(w, x_*, \alpha)| = 0$$

This theorem says that the selection coefficients of the finite problems converge to the selection coefficient of the continuum problem. The full justification of our large population analysis is, however, obtained from Theorem 4.2 which we repeat at this point for the reader's convenience.

Theorem B.2.

$$\lim_{N \to \infty} \lim_{\eta \to 0} |\frac{\eta}{N} \log \frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)} - S(w, x_*, \alpha)| = 0.$$

B.1 Proof of Theorem B.1

We will show that $\gamma_A^N(w, x_*, \alpha) \to \gamma_A(w, x_*, \alpha)$ and $\gamma_B^N(w, x_*, \alpha) \to \gamma_B(w, x_*, \alpha)$ as $N \to \infty$. It suffices to show that $\gamma_A^N(w, x_*, \alpha) \to \gamma_A(w, x_*, \alpha)$ since the same argument can be used to establish existence of the other limit. This will prove Theorem B.1. We proceed in several steps.

Step 1: Let $\bar{\phi}$ denote an optimal control path, defined on the time domain $[0, T]$. By sampling points from the curve at the time points $0 < t_1^N < t_2^N \ldots < t_{T^N}^N$, where $t_k^N := k/N$ and $T^N := \lceil NT \rceil$, we obtain a finite sequence of points

$$\bar{\phi}^N := (\bar{\phi}(0), \bar{\phi}(t_1^N), \ldots, \bar{\phi}(t_{T^N}^N)).$$

For $p \in \mathcal{P}$ and $z \in \{0, 1, \ldots, N^p\}$ let us introduce the family of intervals

$$I_z^{p,N} := (\bar{\phi}^p)^{-1}[z/N, (z+1)/N), 0 \leq z \leq N^p - 1$$
$$I_{N^p}^{p,N} := (\bar{\phi}^p)^{-1}(\{m^p\}).$$

[21] Note that $V \subset U$, so that this cost index is indeed an extension of the discrete cost measure 4.1.

If $t \in I_z^{p,N}$ then $\frac{z}{N} \leq \bar{\phi}^p(t) < \frac{z+1}{N}$. Define the step function

$$\psi^{p,N}(t) := \sum_{z=0}^{N^p-1} \frac{z}{N} \mathbb{1}_{I_z^{p,N}}(t) + m^p \mathbb{1}_{I_{N^p}^{p,N}}(t) \quad (p \in \mathcal{P}, N \geq N_0) \tag{B.2}$$

Evaluating this function at the time points $t_0^N, \ldots, t_{T^N-1}^N$ generates a finite sequence of (not necessarily distinct) points on the grid X^N, $\psi^N(t_k^N), 0 \leq k < T^N$. Geometrically, what we achieve with this construction is a lower approximation of the curve $\bar{\phi}$ with a step function in the (x^1, x^2)-plane. At the terminal point T^N we complete the sequence by taking $x_{T^N}^N$ the closest point to $\bar{\phi}(T^N)$ which lies in the set $\partial\mathsf{X}_B^N$, i.e.

$$\psi^N(T^N) = \arg\min\{||\bar{\phi}(T^N) - x^N|| : x^N \in \partial\mathsf{X}_B^N\}.$$

The so generated sequence is denoted, with an abuse of notation, as

$$\psi^N = (\psi^N(t_0^N), \ldots, \psi^N(t_{T^N})).$$

By construction we have that $\psi^N(t_0^N) = B \equiv (0,0)$. Moreover, $\psi^N(t) \leq \bar{\phi}(t)$ for all $t \in [0,T)$ and on this interval we have even $\psi^N \uparrow \bar{\phi}$ uniformly. To establish uniform convergence on the whole interval $[0,T]$, observe that $T^N \downarrow T$ and that $\psi^N(T^N) \to \bar{\phi}(T)$ as $N \to \infty$. The last point follows from the fact that the closest point projection is Lipschitz. Hence,

$$\lim_{N\to\infty} \max_{0\leq t\leq T} ||\psi^N(t) - \bar{\phi}(t)|| = 0. \tag{B.3}$$

For the two sequences $\bar{\phi}^N$ and ψ^N we can evaluate the action functional as

$$L_{T^N}(\bar{\phi}^N) = \sum_{k=0}^{T^N-1} \left\langle \kappa(\bar{\phi}(t_k^N), \bar{\phi}(t_{k+1}^N) - \bar{\phi}(t_k^N) \right\rangle,$$
$$L_{T^N}(\psi^N) = \sum_{k=0}^{T^N-1} \left\langle \kappa(\psi^N(t_k^N)), \psi^N(t_{k+1}^N) - \psi^N(t_k^N) \right\rangle.$$

Choose N sufficiently large, say $N \geq N_0$, so that for all $N \geq N_0$ the jumps of the path ψ^N satisfy the constraints imposed on the admissible paths for the dynamic program. Then ψ^N is an admissible path for the dynamic program (A.4). Therefore it is true that

$$L_{T^N}(\psi^N) \geq \gamma_A^N(w, x_*, \alpha) \quad \forall N \geq N_0. \tag{B.4}$$

Step 2: Given an arbitrary path $\psi : [0,T] \to \mathsf{X}$ we denote its variation as

$$\mathcal{V}(\psi) := \sup\{\sum_{k=0}^{n-1} ||\psi(t_{k+1}^N) - \psi(t_k^N)|| : 0 = t_0 < t_1 < \cdots < t_n = T\},$$

where the supremum is taken with respect to all partitions of $[0,T]$. A curve ψ is of bounded variation, or rectifiable, if $\mathcal{V}(\psi) < \infty$. It is easy to see that all our paths are of bounded variation, i.e. rectifiable. For every $N \geq N_0$ the action functional $L_{T^N}(\bar{\phi}^N)$ is a Riemann-Stieltjes sum with respect to the curve $\bar{\phi}$, and the equidistant covering of $[0, T]$ introduced above. Since $\bar{\phi}$ is absolutely continuous, it is of bounded variation, and hence rectifiable. Since κ is uniformly continuous, the definition of line integrals as a sequence of Riemann-Stieltjes sums gives us the following immediate fact:

$$\lim_{N\to\infty} L_{T^N}(\bar{\phi}^N) = \int_0^T c(\bar{\phi}(t), \bar{\phi}'(t))dt. \tag{B.5}$$

Step 3: Eq. (B.5) will be used to prove that

$$\lim_{N\to\infty} |L_{T^N}(\psi^N) - L_{T^N}(\bar{\phi}^N)| = 0.$$

By an iterated application of the triangle inequality we get

$$\begin{aligned}
&|L_{T^N}(\psi^N) - L_{T^N}(\bar{\phi}^N)| \\
&\quad \leq \sum_{k=0}^{T^N-1} |\left\langle \kappa(\psi^N(t_k^N)), \psi^N(t_{k+1}^N) - \psi^N(t_k^N)\right\rangle - \left\langle \kappa(\bar{\phi}(t_k^N)), \bar{\phi}^N(t_{k+1}^N) - \bar{\phi}^N(t_k^N)\right\rangle|.
\end{aligned}$$

Now, observe that we can bound each individual term in this sum as follows:

$$\begin{aligned}
&|\left\langle \kappa(\psi^N(t_k^N)), \psi^N(t_{k+1}^N) - \psi^N(t_k^N)\right\rangle - \left\langle \kappa(\bar{\phi}(t_k^N)), \bar{\phi}(t_{k+1}^N) - \bar{\phi}(t_k^N)\right\rangle| \\
&= |\left\langle \kappa(\psi^N(t_k^N)) - \kappa(\bar{\phi}(t_k^N)), \psi^N(t_{k+1}^N) - \psi^N(t_k^N)\right\rangle \\
&- \left\langle \kappa(\bar{\phi}(t_k^N)), [(\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N)) - (\bar{\phi}(t_k^N) - \psi^N(t_k^N))]\right\rangle| \\
&\leq |\left\langle \kappa(\psi^N(t_k^N)) - \kappa(\bar{\phi}(t_k^N)), \psi^N(t_{k+1}^N) - \psi^N(t_k^N)\right\rangle| \\
&+ |\left\langle \kappa(\bar{\phi}(t_k^N)), [(\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N)) - (\bar{\phi}(t_k^N) - \psi^N(t_k^N))]\right\rangle| \\
&\leq ||\kappa(\psi^N(t_k^N) - \kappa(\bar{\phi}(t_k^N)|| \cdot ||\psi^N(t_{k+1}^N) - \psi^N(t_k^N)|| \\
&+ |\left\langle \kappa(0), [(\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N)) - (\bar{\phi}(t_k^N) - \psi^N(t_k^N))]\right\rangle|
\end{aligned}$$

In this derivation we have used in the first bound the triangle inequality, and in the second bound the Cauchy-Schwarz inequality, together with the fact that the largest cost always occur at the beginning of the exit path, i.e. $\kappa(0) \geq \kappa(x)$ for all $x \in \mathsf{X}_B$. Setting $K := \max\{\kappa^1(0), \kappa^2(0)\}$, we observe that

$$\begin{aligned}
&|\left\langle \kappa(0), (\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N) - (\bar{\phi}(t_k^N) - \psi^N(t_k^N))\right\rangle| \\
&\leq K \cdot ||(\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N) - (\bar{\phi}(t_k^N) - \psi^N(t_k^N))|| \\
&\leq K \left(||(\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N)|| + ||(\bar{\phi}(t_k^N) - \psi^N(t_k^N))||\right).
\end{aligned}$$

This bound allows us to conclude that

$$\begin{aligned}
&\sum_k |\left\langle \kappa(0), [(\bar{\phi}(t_{k+1}^N) - \psi^N(t_{k+1}^N)) - (\bar{\phi}(t_k^N) - \psi^N(t_k^N))]\right\rangle| \\
&\qquad \leq 2K \max_{0\leq t\leq T} ||\bar{\phi}(t) - \psi^N(t)||.
\end{aligned}$$

By eq. (B.3) we find for every $\epsilon > 0$ an index $N_0 = N(\epsilon)$ such that for all $N \geq N_0$

$$\max_{0\leq t\leq T} ||\bar{\phi}(t) - \psi^N(t)|| < \epsilon.$$

This fact, and the basic estimate above, allows us to bound the differences in the action functionals for $N \geq N_0$ as

$$|L_{T^N}(\psi^N) - L_{T^N}(\bar{\phi}^N)| \leq \max_{0\leq t\leq T} ||\kappa(\psi^N(t)) - \kappa(\bar{\phi}(t))||\mathcal{V}(\psi^N) + \epsilon.$$

As $N \to \infty$ the step function $\psi^N(\cdot)$ converges uniformly to the optimal control path. The function $\kappa : \mathsf{X} \to \mathbb{R}^2$ is uniformly continuous and X compact. Hence, we can find for every $\epsilon > 0$ a population size N_1 such that for all $N \geq N_1$

$$\max_{0\leq t\leq T} ||\kappa(\bar{\phi}(t)) - \kappa(\psi^N(t))|| < \epsilon.$$

Hence, choosing $N_2 = \max\{N_0, N_1\}$ gives us

$$|L_{T^N}(\psi^N) - L_{T^N}(\bar{\phi}^N)| < \epsilon$$

for all $N \geq N_2$. We conclude that

$$\lim_{N\to\infty} |L_{T^N}(\psi^N) - L_{T^N}(\bar{\phi}^N)| = 0.$$

Together with eq. (B.5) this shows that

$$\lim_{N\to\infty} L_{T^N}(\bar{x}^N) = \gamma_A(w, x_*, \alpha). \tag{B.6}$$

Step 4: We claim that

$$\gamma_A^N(w, x_*, \alpha) \geq \gamma_A(w, x_*, \alpha) \quad \forall N \geq N_0. \tag{B.7}$$

To prove this, let $\bar{x}^N = (\bar{x}_0^N, \ldots, \bar{x}_{K^N}^N)$ denote a solution of the dynamic program. Given this data set let us again introduce the time points t_k^N, so that $0 < t_1^N < t_2^N \ldots < t_{K^N}^N \equiv T^N = \frac{1}{N}K^N$. Then define the polygonal interpolation $\tilde{\psi}^N : [0, T^N] \to \mathsf{X}$ as

$$\begin{aligned}
&\tilde{\psi}^N(0) := (0, 0),\\
&\tilde{\psi}^N(t) := \bar{x}_k^N + (Nt - k)(\bar{x}_{k+1}^N - \bar{x}_k^N), \quad t \in [k/N, (k+1)/N).
\end{aligned}$$

On each interval $[k/N, (k+1)/N)$ the curve $\tilde{\psi}^N$ is continuously differentiable, where derivative at the boundary has to be understood as the adequate one-sided derivative. Then we observe that

$$\begin{aligned}
L_{T^N}(\tilde{\psi}^N) &= \sum_{k=0}^{K^N-1} \int_{k/N}^{(k+1)/N} \left\langle \kappa(\tilde{\psi}^N(t_k^N)), (\tilde{\psi}^N)'(t) \right\rangle dt\\
&\leq \sum_{k=0}^{K^N-1} \frac{1}{N} \left\langle \kappa(\bar{x}_k^N), N(\bar{x}_{k+1}^N - \bar{x}_k^N) \right\rangle\\
&= \gamma_A^N(w, x_*, \alpha).
\end{aligned}$$

The inequality in the second line follows from the fact along the path out of the stability set of B, costs are non-increasing, and therefore $\kappa(\bar{x}_k^N)$ is the largest cost point on the line connecting $\bar{x}_k^N$ with $\bar{x}_{k+1}^N$. Now, observe that each polygonal path $\tilde{\psi}^N(\cdot)$, generated by the data provided by the solution of the dynamic program (A.4), is an admissible curve for the optimal control problem (4.7). This follows from the fact that $\tilde{\psi}^N(0) = B = (0,0)$, $\tilde{\psi}^N(T^N) \notin \mathsf{X}_B$, and $\tilde{\psi}^N(\cdot)$ is piecewise continuously differentiable with admissible tangent vectors. Hence, it is true that

$$\gamma_A^N(w, x_*, \alpha) \geq \gamma_A(w, x_*, \alpha) \quad \forall N \geq N_0. \tag{B.8}$$

Step 5: We have shown in Step 1 and Eq. (B.7) that

$$L_{T^N}(\psi^N) \geq \gamma_A^N(w, x_*, \alpha) \geq \gamma_A(w, x_*, \alpha)$$

for all $N \geq N_0$. Now, from Step 3 we know that $L_{T^N}(\psi^N) \to \gamma_A(w, x_*)$. Hence, it follows that $\gamma_A^N(w, x_*) \to \gamma_A(w, x_*)$ as well. This completes the proof.

B.2 Proof of Theorem 4.2

Note that for every $N \geq N_0$ we have that

$$\begin{aligned}|\frac{\eta}{N}\log\frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)} - S(w, x_*, \alpha)| \leq |\frac{\eta}{N}\log\frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)} - S^N(w, x_*, \alpha)|\\ +|S^N(w, x_*, \alpha) - S(w, x_*, \alpha)|\end{aligned}$$

Taking first $\eta \to 0$ we get from Proposition 4.1 that

$$|\frac{\eta}{N}\log\frac{\mu^{N,\eta}(B)}{\mu^{N,\eta}(A)} - S^N(w, x_*, \alpha)| \to 0$$

and then we take $N \to \infty$ and cite Theorem B.1. This completes the proof of Theorem 4.2.

Appendix C Solutions to the optimal control problems

We now present a complete solution to the optimal control problem (4.11), and *a-fortiori* of problem (4.12). Recall that the Hamiltonian associated with this problem is given by

$$H(x(t), u(t), \lambda(t)) = I(x(t), \lambda(t)) + u(t)G(x(t), \lambda(t)),$$

where

$$\begin{aligned}I(x, \lambda) &:= \lambda^2 - \lambda_0 w(\alpha(x_*^1 - x)),\\ G(x, \lambda) &: = \lambda_0\left(w[\alpha(x_*^1 - x^1)] - w(x_*^2 - x^2)\right) + \lambda^1 - \lambda^2.\end{aligned}$$

Necessary conditions for a solution $(\bar{x}, \bar{u}, T)$ are (see e.g. Seierstadt and Sydsaeter, 1987, pp.83 and pp. 143):

1. $(\lambda_0, \lambda_1(t), \lambda_2(t)) \neq (0,0,0)$, $\lambda_0 \in \{0,1\}$,
2. $\bar{u}(t) \in \arg\max_{u \in [0,1]} H(\bar{x}(t), u, \bar{\lambda}(t))$ for almost all $t \in [0,T]$;
3. $\frac{d}{dt}\bar{\lambda}^p(t) = -\frac{\partial H}{\partial x^p}(\bar{x}(t), \bar{u}(t), \lambda(t))$ whenever $\bar{u}$ is continuous,

4. $\bar{\lambda}^p(T) \leq 0$, with equality if $x^p(T) < x_*^p, p = 1, 2$;
5. $H(\bar{x}(T), \bar{\lambda}(T), \bar{u}(T)) = 0$.

The situation $\lambda_0 = 0$ is particular since in this case the cost functional would not influence the solution. The following Lemma shows that we can safely ignore this case.

Lemma C.1. $\lambda_0 = 1$ *is always true.*

Proof. For sake of contradiction suppose that $\lambda_0 = 0$. Then the Hamiltonian reduces to $H = u(\lambda^1 - \lambda^2) + \lambda^2$. By the maximum principle, $\bar{u} = 1$ if $\lambda^1 > \lambda^2$. Since the Hamiltonian is independent of x, the adjoint functions are constant, i.e. $\lambda^p \equiv \bar{\lambda}^p$. Suppose that $\bar{\lambda}^1 = \bar{\lambda}^2$. Then u can be chosen arbitrarily in $[0, 1]$. Suppose $u \in (0, 1)$. Independently of the choice of the control, the condition $H(x(T), u(T), \lambda(T)) = 0$ implies that $\bar{\lambda}^2 = 0$, a contradiction. Now, without loss of generality we may assume that $\bar{\lambda}^1 < \bar{\lambda}^2$. Then $\bar{u} = 0$ is optimal, and it must be true that $\bar{\lambda}^2 = 0$. However, since $\bar{x}^1(T) < x_*^1$, transversality requires that $\bar{\lambda}^1 = 0$, a contradiction. □

Given this fact we can exclude λ_0 from our set of adjoint functions. We use the notation $\lambda := (\lambda^1, \lambda^2)$ for the adjoint functions. These functions solve the differential equations

$$\dot{\lambda}^1(t) = -\frac{\partial H}{\partial x^1}[x(t), u(t), \lambda] = -(1-u)\alpha w'[\alpha(x_*^1 - x^1)], \tag{C.1}$$

$$\dot{\lambda}^2(t) = -\frac{\partial H}{\partial x^2}[x(t), u(t), \lambda] = -uw'(x_*^2 - x^2). \tag{C.2}$$

Since the Hamiltonian is linear in the control variable we can formulate the Maximum principle as

$$\bar{u}(t) = \begin{cases} 1 & \text{if } G(x(t), \lambda(t)) > 0, \\ 0 & \text{if } G(x(t), \lambda(t)) < 0. \end{cases} \tag{C.3}$$

If $G(x(t), \lambda(t)) \equiv 0$ occurs on a path, the decision rule (C.3) gives us no information where we should direct the path to. If this occurs then we have

$$w[\alpha(x_*^1 - x^1)] - w(x_*^2 - x^2) = \lambda^2 - \lambda^1. \tag{C.4}$$

Taking time derivatives in (C.4) and using (C.1)-(C.2) we obtain the marginal-cost equalization condition

$$w'(x_*^2 - x^2) = \alpha w'[\alpha(x_*^1 - x^1)]. \tag{C.5}$$

This equation determines a one-dimensional manifold in the (x^1, x^2)-plane by the map

$$h(x^1, x^2) := w'(x_*^2 - x^2) - \alpha w'[\alpha(x_*^1 - x^1)] \equiv 0.$$

The set $\Sigma := h^{-1}(0)$ is the *singular surface*, a one-dimensional manifold on which $u \in (0, 1)$ holds. If we assume that κ is C^2, we can compute along a path $x(t) = (x^1(t), x^2(t)) \in \Sigma$

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\,t} h(x^1(t), x^2(t)) = -(1-u(t))w''(x_*^2 - x^2(t)) + \alpha^2 u(t) w''[\alpha(x_*^1 - x^1(t))] \equiv 0 \\ &\Rightarrow u(t) = g(x(t); x_*, \alpha) := \frac{w''(x_*^2 - x^2(t))}{w''(x_*^2 - x^2(t)) + \alpha^2 w''[\alpha(x_*^1 - x^1(t))]} \end{aligned}$$

Suppose a path $x(\cdot)$ hits the target set $\partial\mathsf{X}_B$ while following the defining curve of the manifold Σ. Then transversality requires that at the hitting time T we have $H(x(T), u(T), \lambda(T)) = 0$. Since $u(T) \in (0,1)$, it follows that $I(x(T), \lambda(T)) = G(x(T), \lambda(T)) = 0$. Since

$$I(x(T), \lambda(T)) = 0 \Rightarrow \lambda^2(T) = w[\alpha(x^1_* - x^1(T))],$$

it must be true that $x^1(T) = x^1_*$, as transversality requires $\lambda^2(T) \leq 0$. Combining this with the condition $G(x(T), \lambda(T)) = 0$ shows that also $x^2(T) = x^2_*$ must be true. To summarize:

Observation 2. *(a) If a path $x(\cdot)$ hits the target set $\partial\mathsf{X}_B$ by following the defining curve of Σ, it must hit the mixed equilibrium point (x^1_*, x^2_*).*

(b) If $w(\pi) = \max\{\pi, 0\}$, e.g. the unlikelihood function of the logit choice model, then if $\alpha \neq 1$ no singular control surface appears.

Part (b) of this observation has the important implication that for the logit choice function almost always controls of the "bang-bang"-type must be optimal. The only chance for an interior control appears in the non-generic case when $\alpha = 1$.[22]

The next observation recalls a well-known fact that for autonomous optimal control problems the Hamiltonian is a constant of motion under any solution candidate.

Lemma C.2. *Consider a triple $(x(\cdot), u(\cdot), \lambda(\cdot))$ satisfying the Maximum principle. Then the Hamiltonian is a constant of motion, i.e.*

$$\frac{\mathrm{d}}{\mathrm{d}\,t} H[x(t), u(t), \lambda(t)] = 0 \quad \forall t \in [0, T].$$

Proof. We denote by $H_x = (\frac{\partial H}{\partial x^1}, \frac{\partial H}{\partial x^2})^\top$ the column vector of partial derivatives w.r.t. x. Then, using this notation, we see that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\,t} H[x(t), u(t), \lambda(t)] &= \langle H_x[x(t), u(t), \lambda(t)], \dot{x}(t)\rangle + H_u[x(t), u(t), \lambda(t)]\dot{u}(t) \\
&\quad + \left\langle H_\lambda[x(t), u(t), \lambda(t)], \dot{\lambda}(t)\right\rangle \\
&= \langle H_x[x(t), u(t), \lambda(t)], (\dot{x}(t) - H_\lambda[x(t), u(t), \lambda(t)])\rangle \\
&\quad + H_u[x(t), u(t), \lambda(t)]\dot{u}(t)
\end{aligned}$$

where we have used the adjoint equation $\dot{\lambda}(t) = -H_x[x(t), u(t), \lambda(t)]$ in the second line. Since $H_\lambda[x(t), u(t), \lambda(t)] = (u(t), 1 - u(t))^\top = (\dot{x}^1(t), \dot{x}^2(t))^\top = \dot{x}(t)^\top$, this reduces to

$$\frac{\mathrm{d}}{\mathrm{d}\,t} H[x(t), u(t), \lambda(t)] = H_u[x(t), u(t), \lambda(t)]\dot{u}(t).$$

If $u(t)$ is either 0 or 1 on some interval (t_0, t_1), then $\dot{u}(t) = 0, \forall t \in (t_0, t_1)$, and the claim follows. If $u(t) \in (0,1)$ on some interval (t_0, t_1), then the Maximum principle gives us $H_u[x(t), u(t), \lambda(t)] = 0, \forall t \in (t_0, t_1)$. $\square$

From this observation we obtain the following result:

[22] Genericity refers here to the pure-strategy payoffs in the game.

Corollary C.1. *Every solution candidate* $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{\lambda}(\cdot), T)$ *must satisfy*

$$(\forall t \in [0, T]) : H(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t)) = 0. \tag{C.6}$$

Proof. The terminal point T must be chosen such that $H(\bar{x}(T), \bar{u}(T), \bar{\lambda}(T)) = 0$. Since the Hamiltonian is a constant of motion, the triple $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{\lambda}(\cdot))$ must stay on the same level set of the Hamiltonian for all times. □

Therefore, if a path is on the manifold Σ not only $h(x^1(t), x^2(t)) \equiv 0$ must be true, but also the adjoint functions are completely determined as

$$\lambda^1(t) = w(x_*^2 - x^2(t)),\ \lambda^2(t) = w[\alpha(x_*^1 - x^1(t))]. \tag{C.7}$$

This holds since $I(x(t), \lambda(t)) = G(x(t), \lambda(t)) \equiv 0$. Given all this information we are now ready to verify Propositions 4.3 and 4.4.

C.1 Proof of Proposition 4.3

Suppose there exists an optimal control path with $\bar{u}(t) = 0$ for $t \in [0, \tau)$ for some $\tau \leq T$. This means that $\bar{x}^1(t) = 0$ and $\bar{x}^2(t) = t, t \in [0, \tau)$. Then from Lemma C.2 we know that for all $t \in [0, \tau)$

$$H(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t)) = I(\bar{x}(t), \bar{\lambda}(t)) = 0 \Rightarrow \bar{\lambda}^2(t) \equiv w(\alpha x_*^1).$$

From (C.1) we get

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\bar{\lambda}(t) = -\alpha w'(\alpha x_*^1) \equiv -a \Rightarrow \bar{\lambda}^1(t) = -at + K^1.$$

The constant K^1 must satisfy the condition

$$G(\bar{x}(0), \bar{\lambda}(0)) = K^1 - w(x_*^2) < 0 \Rightarrow K^1 < w(x_*^2).$$

Additionally, at time $t = \tau$ there should be a change in the control. Hence $G(\bar{x}(\tau), \bar{\lambda}(\tau)) = -a\tau + K^1 - w(x_*^2 - \tau) = 0$, form which it follows that

$$\tau = \frac{K^1 - w(x_*^2 - \tau)}{a}. \tag{C.8}$$

At $t = \tau$ the path $\bar{x}(\cdot)$ hits the manifold Σ. Hence, for all $t \geq \tau$ the state variables satisfy the relations

$$\bar{x}^1(t) = \int_\tau^t \bar{u}(s)\,\mathrm{d}\,s,\ \bar{x}^2(t) = t - \bar{x}^1(t)$$

where $\bar{u}(t) = g(\bar{x}(t); x_*, \alpha)$ as defined in equation (4.14). By (C.5) we get the condition

$$h(\bar{x}^1(\tau), \bar{x}^2(\tau)) = 0 \Rightarrow w'(x_*^2 - \tau) = \alpha w'(\alpha x_*^1) = a.$$

Under strict convexity of the waste function w, its first derivative w' is C^1 and monotonically increasing, and therefore possesses a continuous inverse function $(w')^{-1}$, which is also monotonically increasing. Then we obtain

$$\tau = x_*^2 - (w')^{-1}(a), \tag{C.9}$$

provided that this expression is nonnegative. This is the case if and only if $w'(x_*^2) \geq a = \alpha w'(\alpha x_*^1)$.[23] Assuming that this relation holds, we can use (C.8) to determine the constant K^1 as

$$K^1 = w(x_*^2 - \tau) + a\tau. \tag{C.10}$$

From time τ onwards the path $\bar{x}(\cdot)$ is on the singular surface Σ, where the adjoint functions are given by (C.7). From Observation 2 we know that $\bar{x}(t) \to (x_*^1, x_*^2)$ as $t \to T$. Since $\bar{x}^1(t) + \bar{x}^2(t) = t$, $\forall t$, it immediately follows that $T = x_*^1 + x_*^2$. It remains to show that the constructed path satisfies the Maximum principle on the interval $[0, \tau)$. Note that for all $t \in [0, \tau)$

$$\begin{aligned} G(\bar{x}(t), \bar{\lambda}(t)) &= \bar{\lambda}^1(t) - w(x_*^2 - t) \\ &= a(\tau - t) + w(x_*^2 - \tau) - w(x_*^2 - t). \end{aligned}$$

At $t = 0$ this reduces to

$$G(\bar{x}(0), \bar{\lambda}(0)) = a\tau + w(x_*^2 - \tau) - w(x_*^2).$$

By the mean-value theorem there exists a $\xi \in (x_*^2 - \tau, x_*^2)$ such that

$$w(x_*^2) - w(x_*^2 - \tau) = w'(\xi)\tau.$$

Since $x_*^2 - \tau = (w')^{-1}(a)$ and w' is monotonically increasing, it follows that $w'(\xi) > a$. Hence

$$G(\bar{x}(0), \bar{\lambda}(0)) = a\tau - w'(\xi)\tau < a\tau - a\tau = 0.$$

Additionally, we observe that

$$\frac{\mathrm{d}}{\mathrm{d}\,t} G(\bar{x}(t), \bar{\lambda}(t)) = w'(x_*^2 - t) - a > w'(x_*^2 - \tau) - a = w'\left((w')^{-1}(a)\right) - a = 0.$$

Consequently G is initially negative and monotonically increasing on $[0, \tau)$, with unique root at $t = \tau$. Setting $a \equiv a(w; x_*, \alpha) \leq b(w; x_*, \alpha) = w'(x_*^2)$ and $\tau = \tau(w; x_*, \alpha)$, we have verified that the proposed triple $(\bar{x}(\cdot), \bar{u}(\cdot), T)$ satisfies all conditions provided by the Maximum principle. Moreover, the adjoint functions satisfy the transversality conditions. The construction of an optimal control path with $u(t) = 1$ on $[0, \tau)$ is similar. From Lemma C.2 we know that

$$H(\bar{x}(t), \bar{u}(t), \lambda(t)) = 0 \Rightarrow \lambda^1(t) = w(x_*^2).$$

The adjoint equations (C.1)-(C.2) tell us that

$$\dot{\lambda}^1(t) = 0, \ \dot{\lambda}^2(t) = -w'(x_*^2) \equiv -b$$

on $[0, \tau)$. Hence, $\lambda^1(t) \equiv w(x_*^2)$ and $\lambda^2(t) = -bt + K^2$ for all $t \in [0, \tau)$. The regime-switching time τ is defined by the condition

$$G(\bar{x}(\tau), \lambda(\tau)) = 0 \Rightarrow \tau = \frac{K^2 - w[\alpha(x_*^1 - \tau)]}{b}.$$

[23] It is not difficult to verify that in the case where $w'(x_*^2) = \alpha w'(\alpha x_*^1)$ it is true that $\tau = 0$. Hence, in this situation the path directly starts from the singular surface Σ.

At $t = \tau$ the path hits the manifold Σ and we get therefore the additional condition

$$w'(x_*^2) = \alpha w'[\alpha(x_*^1 - \tau)] \Rightarrow \tau = x_*^1 - \frac{1}{\alpha}(w')^{-1}(b/\alpha)$$

where we have exploited the strict convexity of the unlikelihood function. We see that $\tau > 0$ if and only if $b(w; x_*, \alpha) = w'(x_*^2) < \alpha w'(\alpha x_*^1) = a(w; x_*, \alpha)$. The constant K^2 is then given by

$$K^2 = w[\alpha(x_*^1 - \tau)] + b\tau.$$

Hence, the adjoint function on $[0, \tau)$ are $\lambda^1(t) = w(x_*^2)$, and $\lambda^2(t) = b(\tau - t) + w[\alpha(x_*^1 - \tau)]$. From time τ onwards the path follows the defining curve of the manifold Σ, whose shape we already know. It remains to show that the Maximum principle is satisfied on $[0, \tau)$. At $t = 0$ we see that

$$G(\bar{x}(0), \bar{\lambda}(0)) = -b\tau + w(\alpha x_*^1) - w[\alpha(x_*^1 - \tau)] = -b\tau + w'(\xi)\alpha\tau$$

where $\xi \in (\alpha(x_*^1 - \tau), \alpha x_*^1)$, whose existence is guaranteed by the mean-value theorem. From the monotonicity of κ' it follows that $\alpha w' > b$, and consequently

$$G(\bar{x}(0), \bar{\lambda}(0)) > -b\tau + b\tau = 0.$$

Since,

$$\frac{\mathrm{d}}{\mathrm{d}\,t} G(\bar{x}(t), \bar{\lambda}(t)) = b - \alpha w'[\alpha(x_*^1 - \tau)] \leq b - \alpha w'\left((w')^{-1}(b/\alpha)\right) = 0$$

the result follows as above.

Since the Maximum principle gives only necessary conditions for optimal solutions, we don't actually know whether the solutions of Propositions 4.3-4.4 are really optimal. To verify this, denote by $(\bar{x}(\cdot), \bar{u}(\cdot))$ a candidate for a solution identified in Propositions 4.3 and 4.4, respectively. Consider a vector $\lambda(t) = (\lambda^1(t), \lambda^2(t))$ such that

$$\dot{\lambda}^p(t) = -\frac{\partial H}{\partial x^p}[\bar{x}(t), \bar{u}(t), \lambda(t)], \quad p = 1, 2 \tag{C.11}$$

$$\frac{\partial H}{\partial u}[\bar{x}(t), \bar{u}(t), \lambda(t)](\bar{u}(t) - u) \geq 0 \quad \forall u \in [0, 1], \forall t \in [0, T] \tag{C.12}$$

$H(x, u, \lambda(t))$ is concave in (x, u) for all t. By the Mangasarian sufficiency theorem (Seierstadt and Sydsaeter, 1987, pp. 103) $(\bar{x}(\cdot), \bar{u}(\cdot))$ is indeed a solution of the optimal control problem (4.11), respectively (4.12). In particular, we can choose $\lambda(t) = \bar{\lambda}(t)$ to test the conditions. It is easily verified that

$$\frac{\partial H}{\partial u}[\bar{x}(t), \bar{u}(t), \bar{\lambda}(t)](\bar{u}(t) - u) = G(\bar{x}(t), \bar{\lambda}(t))(\bar{u}(t) - u) \geq 0$$

for all $u \in [0, 1]$ and all $t \in [0, T]$. Hence, the pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ satisfies also all sufficient conditions. Hence, we have proved the following result:

Proposition C.1. *The triple $(\bar{x}(\cdot), \bar{u}(\cdot), T)$ identified in Propositions 4.3-4.4 is a solution of the optimal control problem (4.11), respectively (4.12).*

Having verified that the identified solution candidates are indeed optimal solutions, we are ready to present the general formulas for the stochastic potentials of the two strict Nash equilibria. Let us introduce the maps $\hat{c}^1(x^2) := w(|x_*^2 - x^2|), \hat{c}^2(x^1) := w(\alpha|x_*^1 - x^1|)$, and the 1-form

$$\omega := \hat{c}^1 dx^1 + \hat{c}^2 dx^2, \tag{C.13}$$

where dx^1, dx^2 are the differentials associated with the canonical coordinate functions of $\mathbb{R}^2$. The 1-form ω is continuously differentiable on $\mathsf{X} \setminus \{(x_*^1, x_*^2)\}$. Denote by $T_A = x_*^1 + x_*^2$ and $T_B = M - T_A$, the respective hitting times of the optimal control paths identified for problems (4.11) and (4.12). Moreover, let us call $\bar{\phi}_{BA} : [0, T_A] \to \mathsf{X}_B$ and $\bar{\phi}_{AB} : [0, T_B] \to \mathsf{X}_A$ the curves corresponding to the respective optimal control paths of problems (4.11) and (4.12). These curves are continuously differentiable almost everywhere. Consequently, we can formulate the stochastic potentials in the elegant way[24]

$$\gamma_A(w, x_*, \alpha) = \int_{\bar{\phi}_{BA}} \omega := \int_0^{T_A} \left(\hat{c}^1(\bar{\phi}_{BA}^2(t))(\bar{\phi}_{BA}^1)'(t) + \hat{c}^2(\bar{\phi}_{BA}^1(t))(\bar{\phi}_{BA}^2)'(t)\right) \mathrm{d}\, t,$$

and similarly

$$\gamma_B(w, x_*, \alpha) = -\int_{\bar{\phi}_{AB}} \omega := -\int_0^{T_B} \left(\hat{c}^1(\psi_{AB}^2(t))(\bar{\phi}_{AB}^1)'(t) + \hat{c}^2(\bar{\phi}_{AB}^1(t))(\bar{\phi}_{AB}^2)'(t)\right) \mathrm{d}\, t.$$

Propositions 4.3 and 4.4 give us a complete characterization of optimal control paths for all noisy-best response protocols which have strictly convex unlikelihood functions on $(0, \infty)$. As a corollary we obtain equilibrium selection results for the probit choice function.

C.2 Solution for the Logit choice

We proof for each of the three cases that the proposed solution satisfies all the necessary conditions for an optimal solution.

(a) We need to verify that the proposed solution satisfies the necessary conditions given by Maximum principle and the appropriate transversality conditions (see Seierstadt and Sydsaeter, 1987, pp.85):

$$\lambda^1(T) \text{ no condition} \tag{C.14}$$
$$\lambda^2(T) \leq 0 \text{ with equality if } x^2(T) < x_*^2. \tag{C.15}$$

The adjoint functions must satisfy equations (C.1)-(C.2), which are in our case

$$\dot{\lambda}^1 = 0,\ \dot{\lambda}^2(t) = -1 \Rightarrow \lambda^1(t) \equiv A\ , \lambda^2(t) = -t + B,$$

for constants A, B which have to be adequately chosen. The Hamiltonian is given by

$$H(\bar{x}(t), \bar{u}(t), \bar{\lambda}(t)) = \bar{\lambda}^1(t) - x_*^2 = A - x_*^2 = 0$$
$$\Rightarrow A = x_*^2.$$

[24]The minus sign in the definition of $\gamma_B(w, x_*, \alpha)$ comes from the fact that $\bar{\phi}'_{AB} \leq 0$ almost everywhere. We could get rid of the minus sign by integrating over the time-reversed path, so the way of writing the formula is a matter of taste.

The control $\bar{u}$ satisfies the Maximum principle if and only if

$$\begin{aligned} G(\bar{x}(t), \lambda(t)) &= \alpha(x_*^1 - t) - x_*^2 + x_*^2 + t - B \\ &= \alpha x_*^1 - B + t(1-\alpha) > 0 \end{aligned}$$

for all $t \in [0, T)$, with an eventual equality at the terminal point $t = T = x_*^1$. At $t = 0$ this requires $B < \alpha x_*^1$. Transversality requires that $\lambda^2(T) = 0$. Therefore $B = x_*^1 < \alpha x_*^1$ iff $\alpha > 1$. We observe that $\frac{\mathrm{d}}{\mathrm{d}\,t} G(\bar{x}(t), \lambda(t)) = 1 - w < 0$ for $\alpha > 1$, and G vanishes exactly at the point $t = x_*^1$. Hence, $G(\bar{x}(t), \bar{\lambda}(t)) \geq 0$ for all $t \in [0, T]$ with equality only at the terminal point $t = T = x_*^1$. Hence, the triple $(\bar{x}, \bar{u}, T)$ satisfies all conditions provided by the Maximum principle, and is therefore a candidate for an optimum. Since $\frac{\partial H}{\partial u} = G \geq 0$, we observe that

$$G(\bar{x}(t), \bar{\lambda}(t))(\bar{u}(t) - u) \geq 0 \quad \forall u \in [0,1], \forall t \in [0, T].$$

The control region is the unit interval $[0, 1]$, a convex and compact set, and the adjoint functions are continuously differentiable. Moreover $H(x, u, \lambda)$ is concave in (x, u). Thus, the proposed solution satisfies the Mangasarian sufficiency conditions (Seierstadt and Sydsaeter, 1987, Theorem 4, pp. 105) and is therefore a solution to the auxiliary optimal control problem.

(b) As in (a) we need to verify that the proposed solution satisfies the Maximum principle with the appropriate transversality conditions

$$\lambda^1(T) \leq 0, \text{ with equality if } x^1(T) < x_*^1 \tag{C.16}$$

$$\lambda^2(T) \text{ no condition} \tag{C.17}$$

Under $\bar{u}(t) \equiv 0$ the adjoint equations (C.1)-(C.2) read as

$$\begin{aligned} \dot{\lambda}^1(t) &= -\alpha \Rightarrow \lambda^1(t) = -\alpha t + A \\ \dot{\lambda}^2(t) &= 0 \Rightarrow \lambda^2(t) \equiv B \end{aligned}$$

The Hamiltonian is given by

$$\begin{aligned} H(\bar{x}(t), \bar{u}(t), \lambda(t)) &= \lambda^2(t) - \alpha x_*^1 = 0 \\ &\Rightarrow B = \alpha x_*^1 \end{aligned}$$

The proposed control $\bar{u}$ satisfies the Maximum principle iff

$$\begin{aligned} G(\bar{x}(t), \lambda(t)) &= \alpha x_*^1 - (x_*^2 - t) - \alpha t + A - \alpha x_*^1 \\ &= A - x_*^2 + t(1-\alpha) \leq 0 \end{aligned}$$

for all $t \in [0, T]$, with equality eventually at the terminal point $t = T = x_*^2$. At $t = 0$ this requires that $A \leq x_*^2$. Transversality requires that $\lambda^1(T) = -\alpha T + A = 0 \Rightarrow A = \alpha x_*^2 < x_*^2$ iff $\alpha < 1$. Moreover $\frac{\mathrm{d}}{\mathrm{d}\,t} G(\bar{x}(t), \bar{\lambda}(t)) = 1 - \alpha > 0$ iff $\alpha < 1$. Hence G is initially negative and monotonically increasing over time, with unique root at $t = T = x_*^2$. Hence, the proposed solution satisfies the Maximum principle and the adjoint functions satisfy the corresponding transversality conditions. Sufficiency follows again from the Mangasarian sufficiency theorem. We have $\frac{\partial H}{\partial u} = G \leq 0$, with equality only at $t = T$. Thus,

$$G(\bar{x}(t), \bar{\lambda}(t))(\bar{u}(t) - u) \geq 0 \quad \forall u \in [0,1], \forall t \in [0, T].$$

(c) If $\alpha = 1$ the adjoint equations (C.1)-(C.2) are

$$\dot{\lambda}^1(t) = -(1 - u(t)) = -\dot{x}^2(t) \Rightarrow \lambda^1(t) = -x^2(t) + A$$
$$\dot{\lambda}^2(t) = -u(t) = -\dot{x}^1(t) \Rightarrow \lambda^2(t) = -x^1(t) + B$$

We see that

$$G(x(t), \lambda(t)) = x_*^1 - x_*^2 + A - B = const$$

and

$$I(x(t), \lambda(t)) = B - x_*^1.$$

If $u = 0$ then $B = x_*^1$. If $u = 1$ then $A = x_*^2$. If $u \in (0, 1)$ then both must be true. Suppose that $\bar{u} = 0$. Then $G(x(t), \lambda(t)) = -x_*^2 + A \leq 0 \Rightarrow A \leq x_*^2$. Transversality requires that $\lambda^1(T) = 0$, so that $A = x_*^2$. Suppose that $\bar{u} = 1$. Then $G(x(t), \lambda(t)) = B - x_*^1 \geq 0 \Rightarrow B \geq x_*^1$. Again, by transversality, we conclude that $B = x_*^1$. Hence, $\bar{u} \in \{0, 1\}$ together with these adjoint functions satisfies the necessary conditions for an optimal control.

This completes the proof of Proposition 4.7.

We can also show that no optimal control can exhibit a jump discontinuity at some intermediate point $\tau \in [0, T]$. Suppose that $\alpha < 1$ and consider the control

$$u(t) = \begin{cases} 0 & \text{if } t \in [0, \tau], \\ 1 & \text{if } t \in (\tau, T] \end{cases}$$

where $\tau \leq T$. The initial segment of the path $x(\cdot)$ generated by the proposed control is as in part (b) of Proposition 4.7. Hence for all $t \in [0, \tau)$ we have

$$x(t) = (0, t), \lambda^1(t) = -\alpha t + A, \lambda^2(t) = \alpha x_*^1, A \leq x_*^2.$$

At $t = \tau$ there is a jump in the control. In terms of the Maximum principle this implies that

$$G(x(\tau), \lambda(\tau)) = 0 \Rightarrow \tau = \frac{x_*^2 - A}{1 - \alpha} \geq 0.$$

On $(\tau, T]$ the adjoint equations (C.1)-(C.2) read as

$$\dot{\lambda}^1(t) = 0 \Rightarrow \lambda^1(t) \equiv A - \alpha\tau = x_*^2 - \tau$$
$$\dot{\lambda}^2(t) = -1 \Rightarrow \lambda^2(t) = -(t - \tau) + \alpha x_*^1.$$

Some formal manipulations give

$$G(x(t), \lambda(t)) = (t - \tau)(1 - \alpha) \geq 0$$

for all $t \geq \tau$ since $w < 1$ by hypothesis. Since $x(t) \to (x_*^1, \tau)$ a transversality condition for any optimal control problem must be $\lambda^2(T) = 0$, or equivalently $(T - \tau) = \alpha x_*^1$. Since $x(T) = x_*^1$, it must be true that $T - \tau = x_*^1 > \alpha x_*^1$, a contradiction. Moreover we need that $\lambda^1(T) \leq 0$, or equivalently $\tau \geq x_*^2$, which is another contradiction. Hence $\tau = T = x_*^2$ must be true, and we obtain the path of part (b) of Proposition 4.7. Similarly we can show that there can be no optimal control of the form

$$u(t) = \begin{cases} 1 & \text{if } t \in [0, \tau], \\ 0 & \text{if } t \in (\tau, T] \end{cases}$$

for $\tau \leq T$ if $\alpha > 1$.

Finally, observe that if $\alpha > 1$, there can be no path that has $u(t) \equiv 0$. This follows immediately from $G(x(0), \lambda(0)) = A - x_*^2 = \alpha x_*^2 - x_*^2 > 0$, contradiction! In the same way we see that there can be no path with $u(t) \equiv 1$ if $\alpha < 1$. This follows because we know that $G(x(0), \lambda(0)) = \alpha x_*^1 - B = (\alpha - 1)x_*^1 < 0$, contradiction!

Alós-Ferrer, C., Netzer, N., 2010. The logit-response dynamics. Games and Economic Behavior 68, 413–427.

Anderson, S., de Palma, A., Thisse, J., 1992. Discrete Choice theory of Product differentiation. The MIT Press.

Bajari, P., Hong, H., Ryan, S. P., 2010. Identification and estimation of a discrete game of complete information. Econometrica 5, 1529–1568.

Benaïm, M., Weibull, J. W., 2003. Deterministic approximation of stochastic evolution in games. Econometrica 71 3, 873–903.

Bergin, J., Lipman, B., 1996. Evolution with state-dependent mutations. Econometrica 64, 943–956.

Binmore, K., Samuelson, L., 1997. Muddling through: Noisy equilibrium selection. Jounal of Economic Theory 74, 235–265.

Binmore, K., Samuelson, L., Vaughn, R., 1995. Musical chairs: Modeling noisy evolution. Games and Economic Behavior 11, 1–35.

Blume, L. E., 1993. Statistical mechanics of strategic interaction. Games and Economic Behavior 5, 387–424.

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

Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Review of Economic Studies 67, 17–45.

Foster, D., Young, H. P., 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology 38, 219–232.

Freidlin, M., Wentzell, A., 1998. Random Perturbations of Dynamical Systems, 2nd Ed. Springer Verlag, New York.

Fudenberg, D., Levine, D. K., 1998. The theory of learning in games. The MIT Press, Cambridge Massachusetts.

Harsanyi, J., Selten, R., 1988. A general theory of equilibrium selection in games. MIT Press, Cambridge Massachusetts.

Hehenkamp, B., 2003. Equilibrium selection in the two-population kmr model. International Game Theory Review 5, no. 3, 249–262.

Hofbauer, J., Sandholm, W. H., 2002. On the global convergence of stochastic fictitious play. Econometrica 70 (6), 2265–2294.

Hofbauer, J., Sandholm, W. H., 2007. Evolution in games with randomly disturbed payoffs. Journal of Economic Theory 132, 47–69.

Hofbauer, J., Sorger, G., 1999. Perfect foresight and equilibrium selection in symmetric potential games. Journal of Economic Theory 85, 1–23.

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

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

Kifer, Y., 1988. Random perturbations of dynamical systems. Birkhäuser - Progress in Probability and Statistics.

Kifer, Y., 1990. A discrete-time version of the Wentzell-Freidlin theory. The Annals of Probability 18, no. 4, 1676–1692.

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

Matsui, A., Matsuyama, K.,1995. An Approach to Equilibrium Selection. Journal of Economic Theory 65, 415–434.

McFadden, D., 1981. Econometric Models of Probabilistic Choice. in Structural Analysis of Discrete Data with Econometric Applications, Cambridge: MIT Press, pp. 198–272.

Morris, S., Rob, R., Shin, H. S., 1995. p-dominance and belief potential. Econometrica 63 (1), 145–157.

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

Oechssler, J., Riedel, F., 2001. Evolutionary dynamics on infinite strategy spaces. Economic Theory 17, 141–162.

Okada, D., Tercieux, O., 2009. Log-linear dynamics and local potential, mimeo.

Sandholm, W. H., 2010a. Population games and Evolutionary dynamics. The MIT Press.

Sandholm, W. H., 2010b. Orders of limits for stationary distributions, stochastic dominance, and stochastic stability. Theoretical Economics 5, 1–26.

Seierstadt, A., Sydsaeter, K., 1987. Optimal control theory with economic applications. Elsevier Science, Advanced Textbooks in Economics.

Stroock, D., 2005. An Introduction to Markov Processes. Graduate Texts in Mathematics 230, Springer Berlin Heidelberg New York.

Young, H., 1993. The evolution of conventions. Econometrica 61 (1), 57–84.

Young, H., 1998. Individual Strategy and Social Structure - An Evolutionary Theory of Institutions. Princeton University Press.