Exit from Equilibrium in Coordination Games under Probit choice*

Srinivas Arigapudi†

September 8, 2018

Abstract

We consider a model of stochastic evolution under the probit choice rule. In the small noise double limit, where first the noise level in agents' decisions is taken to zero, and then the population size to infinity, escape from and transitions between equilibria can be described in terms of solutions to continuous optimal control problems. We use results from optimal control theory to solve the exit cost problem which is used to assess the expected time until the evolutionary process leaves the basin of attraction of a stable equilibrium in a class of three-strategy coordination games.

1. Introduction

Evolutionary game theory studies the behavior of strategically interacting agents whose decisions are based on simple myopic rules. Together, a game, a decision rule, and a population size define a stochastic aggregate behavior process on the set of population states. In general, agents play strategies which are best responses to the current population state. However sometimes they end up choosing strategies that are suboptimal due to noise in the underlying model. Over short to moderate time spans, the process typically settles near a Nash equilibrium of the game. But over longer time spans, breakdown of and transitions between equilibria are inevitable, with some occurring more readily than others. Most work in stochastic stability analysis follow the seminal papers of Kandori et al. (1993) and Young (1993). They model the noise using best response with mutations (BRM) in which the probability of a suboptimal choice is independent of its

*I am greatly indebted to Professor Bill Sandholm for invaluable help on this project. Financial support from Army Research Office Grant MSN201957 and Summer Research Fellowship (2018) from Department of Economics, University of Wisconsin-Madison is gratefully acknowledged.

†Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706. e-mail: arigapudi@ssc.wisc.edu; website: sites.google.com/view/srinivasarigapudi/home.

payoff consequences. This model eases the analysis, as the difficulty of transiting from one equilibrium to another can be determined by counting the number of mutations needed for the transition to occur.

In most applications however, it is more realistic to assume that costly mistakes are less likely to happen, i.e., it is more reasonable to assume that the errors in agents choices are payoff dependent as in the logit model of Blume (1993, 2003) and the probit model of Myatt and Wallace (2003) and Dokumacı and Sandholm (2011). When mistake probabilities are payoff-dependent, the probability of a transition between equilibria becomes more difficult to assess, depending not only on the number of suboptimal choices required, but also on the unlikelihood of each such choice. As a consequence, general results on transitions between equilibria and stochastic stability are only available for two-strategy games using birth-death chain methods.

In this paper, we use dynamic programming methods to characterize equilibrium breakdown in a class of three strategy coordination games under the probit choice rule. The importance of these questions has been recognized in macroeconomics, see Williams (2001), Cho et al. (2002). We consider the *exit problem*, which is used to assess the expected time until the evolutionary process leaves the basin of attraction of a stable equilibrium, and determine the likely exit path. Here, the cost of a path is the sum of unlikelihoods associated with the changes in strategy along the path where the unlikelihood of a strategy is the exponential rate of decay of the probability of choosing it as the noise approaches zero. In the class of games we study here, Sandholm and Staudigl (2016) (henceforth SS16) solved the exit problem when agents follow the logit choice rule. They show that the likely exit path proceeds along the boundary of the simplex, escaping the basin of attraction through a boundary mixed equilibrium (see Figure 2). The logit choice rule has a very simple piecewise linear form for the unlikelihood function (see (2.3)), which makes the computation of path costs relatively straightforward. However, the unlikelihood functions are more complicated under the probit choice. Depending on the population state, the unlikelihood function will take different forms (see Lemma 2.3). Therefore, finding the optimal exit paths is more involved under the probit choice framework.

We solve the exit problem when agents follow the probit choice rule. We show that the solution is qualitatively similar to the logit choice in some cases, with the optimal exit path dividing the initial basin of attraction of the equilibrium into two regions (see Figure 2). We call this the *standard* case. We also find that in certain cases the optimal exit paths divides the initial basin of attraction of the equilibrium into three regions as shown in Figures 3 and 4. We find that in these cases there is a region where the optimal exit path heads back in the direction of the initial equilibrium. We call this the *retreating* case.

The retreating case can be further divided into two subcases: one where the optimal exit path has *non-binding* state constraints as shown in Figure 3, another where the optimal exit path has *binding* state constraints as shown in Figure 4.

The paper proceeds as follows: Section 2 introduces our class of stochastic evolutionary processes. We formally define the unlikelihood function and describe it under the two common noisy best response protocols, *logit choice* and *probit choice*. In Section 3, we set up the *exit cost* problem in the small noise double limit. Section 4.1 provides definitions and introduces notation which will be used in our main analysis for working with symmetric normal form games. Section 4.2 provides outline of analysis and gives an overview of the main results. In Section 5.1, we provide the statement of verification theorem which provides a sufficient condition for a value function to be a solution in a certain class of optimal control problems which includes the exit cost problem. In Section 5.2, we compute the costs of certain direct paths which will later be used to guess the form of the optimal value function in Section 5.3. In Section 6, we verify the optimal value function for the different cases using the verification theorem. Section 7 concludes.

2. Model

2.1 Population Games

We consider games in which agents from a population of size N choose strategies from the common finite strategy set $S = \{1, 2, \cdots, n\}$. The population's aggregate behavior is described by a *population state* x, an element of the simplex $X = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$ (Our main analysis later will focus on cases with three strategies i.e., $n = 3$). The standard basis vector $e_i \in X \subset \mathbb{R}^n$ represents the *pure population state* at which all agents play strategy i. At population state $x = (x_1, x_2, \cdots, x_n)'$, x_i represents the fraction of agents playing strategy i. We suppose that agents are randomly matched to play the symmetric normal form game $A \in \mathbb{R}^{n\times n}$. The $(i, j)^{th}$ entry A_{ij} is the payoff a player obtains when he chooses strategy i and his opponent chooses strategy j. The expected payoff to strategy i at population state x is described by $F_i(x) = \sum_{j=1}^n A_{ij}x_j$. In matrix notation, we have $F(x) = Ax$.[1]

We suppose that the symmetric normal form game A is a *coordination game*: $A_{ii} > A_{ji}$ for all $i, j \in S$, $i \neq j$. This implies that if one's match partner plays i, one is best off playing i oneself. We also suppose that A has *marginal bandwagon property* of Kandori and Rob (1998): $A_{ii} - A_{ik} > A_{ji} - A_{jk}$ for all $i, j, k \in S$ with $i \notin \{j, k\}$. In words, the above condition requires that when some agent switches to strategy i from any other strategy k, current

[1]We will interchangeably use $\pi(x)$ to denote the payoff vector i.e., $\pi(x) = F(x)$.

strategy i players benefit most. Following SS16, we call three strategy coordination games that satisfy marginal bandwagon property and that admit an interior equilibrium, *simple three-strategy coordination games*. These classes of games are large enough to allow some variety in analysis, but small enough that the analysis remains manageable as pointed in SS16.

2.2 Noisy best response protocols and unlikelihood function

In our model of stochastic evolution, agents occasionally receive opportunities to switch strategies. Upon receiving a revision opportunity, an agent selects a strategy by employing a *noisy best response protocol* $\sigma^\eta : \mathbb{R}^n \to \text{int}(X)$ with *noise level* $\eta > 0$, a function that maps vectors of payoffs to vectors of probabilities of choosing each strategy.

For a noisy best response protocol σ^η, the *unlikelihood function* Υ_j represents the exponential rate of decay of the probability that strategy j is chosen as η approaches zero i.e.,

$$\Upsilon_j(\pi) = -\lim_{\eta\to 0} \eta \log \sigma^\eta_j(\pi) \tag{2.1}$$

Example 2.1. Logit choice. The *logit choice protocol* with noise level η, introduced to evolutionary game theory by Blume (1993), is defined by

$$\sigma^\eta_j(\pi) = \frac{\exp(\eta^{-1}\pi_j)}{\sum_{k\in S}\exp(\eta^{-1}\pi_k)}. \tag{2.2}$$

It is easy to verify that this protocol has piecewise linear unlikelihood function given by

$$\Upsilon_j(\pi) = \max_{k\in S} \pi_k - \pi_j \tag{2.3}$$

◆

Example 2.2. Probit choice. The *probit choice protocol* is an additive random utility model in which the payoff vector π is perturbed by adding the sample average $\bar{\epsilon}^m$ of an i.i.d. sequence $\{\epsilon^l\}_{l=1}^m$ of independent standard normal random variables. Writing η for $\frac{1}{m}$, we obtain the protocol

$$\sigma^\eta_j(\pi) = \mathbb{P}\left(j \in \underset{k\in S}{\text{argmax}}(\pi_k + \bar{\epsilon}^m_k)\right) \tag{2.4}$$

Using theory of large deviations, Dokumacı and Sandholm (2011) derive the unlikelihood function when agents follow probit choice. Stating their result needs some notation which we provide in Appendix A. Using their result, we can express the unlikelihood of a state in the best response region of strategy i in a three strategy game A, as follows:

Lemma 2.3. *Let π be a payoff vector with $\pi_i \geq \max\{\pi_j, \pi_k\}$. Then the unlikelihood function Υ under the probit choice rule is given by $\Upsilon_i(\pi) = 0$*

$$\Upsilon_j(\pi) = \begin{cases} \frac{(\pi_i - \pi_j)^2}{4} & \text{if } \pi_k \leq \frac{\pi_i + \pi_j}{2} \\ \frac{(\pi_i - \pi_j)^2 + (\pi_k - \pi_j)^2 + (\pi_i - \pi_k)^2}{6} & \text{if } \pi_k \geq \frac{\pi_i + \pi_j}{2} \end{cases}$$

$$\Upsilon_k(\pi) = \begin{cases} \frac{(\pi_i - \pi_k)^2}{4} & \text{if } \pi_j \leq \frac{\pi_i + \pi_k}{2} \\ \frac{(\pi_i - \pi_j)^2 + (\pi_k - \pi_j)^2 + (\pi_i - \pi_k)^2}{6} & \text{if } \pi_j \geq \frac{\pi_i + \pi_k}{2} \end{cases}$$

Proof. See Appendix A. □

For a base payoff vector $\pi = [\pi_i \;\; \pi_j \;\; \pi_k]$ in the best response response of strategy i, we have $\pi_i \geq \max\{\pi_j, \pi_k\}$. The shock vector $\bar{\epsilon} = [\bar{\epsilon}_i \;\; \bar{\epsilon}_j \;\; \bar{\epsilon}_k]$ which makes strategy j better than strategies i and k should satisfy the equations $\pi_j + \bar{\epsilon}_j \geq \pi_i + \bar{\epsilon}_i$ and $\pi_j + \bar{\epsilon}_j \geq \pi_k + \bar{\epsilon}_k$ i.e.,

$$\bar{\epsilon}_j - \bar{\epsilon}_i \geq \pi_i - \pi_j \text{ and } \bar{\epsilon}_j - \bar{\epsilon}_k \geq \pi_k - \pi_j \tag{2.5}$$

If either $\pi_j \geq \pi_k$ or ($\pi_k \geq \pi_j$ and $\pi_k \leq \frac{\pi_i + \pi_j}{2}$) then large deviation analysis shows that the least unlikely way to satisfy (2.5) is to have a zero shock to action k and the shocks to actions i and j "share the burden equally" i.e., $\bar{\epsilon}_k = 0$, $\bar{\epsilon}_j = \overline{\pi}^{\{i,j\}} - \pi_j$ and $\bar{\epsilon}_i = \overline{\pi}^{\{i,j\}} - \pi_i$, where $\overline{\pi}^{\{i,j\}} = \frac{\pi_i + \pi_j}{2}$.

If $\pi_k \geq \frac{\pi_i + \pi_j}{2}$ then large deviation analysis shows that the least unlikely way to satisfy (2.5) is to have the shocks to all actions "share the burden equally" i.e., $\bar{\epsilon}_l = \overline{\pi} - \pi_l$, where $l = \{i, j, k\}$ and $\overline{\pi} = \frac{\pi_i + \pi_j + \pi_k}{3}$. ◆

3. The Small Noise Double Limit

3.1 Induced Markov Chain

A population game F, a noisy best response protocol σ^η and a population size N generate a Markov chain $\mathbf{X}^{N,\eta}$ on the set of population states X. For coordination games, over short to medium time scales, $\mathbf{X}^{N,\eta}$ typically converges to an equilibrium of the underlying game F. Over longer periods, runs of suboptimal choices occasionally occur, leading to transitions between the equilibria. In the small noise double limit (first $\eta \to 0$, then $N \to \infty$), SS16 characterize these transitions as solutions to control problems whose running costs are obtained by composing the unlikelihood function with the payoff function, as we describe next.

3.2 Costs of continuous paths

Let $\langle .,. \rangle$ denote the standard inner product on $\mathbb{R}^n$. Let $\phi : [0, T] \to X$ be absolutely continuous and *non-pausing*, meaning that $|\dot{\phi}_t| \neq 0$ for almost all $t \in [0, T]$. In the small noise double limit, the cost of the continuous path ϕ_t in the small noise double limit derived in SS16 is given by

$$c(\phi) = \int_0^T \langle \Upsilon(F(\phi_t)), [\dot{\phi}_t]_+ \rangle dt. \tag{3.1}$$

The intuition for (3.1) is as follows: The cost of a discrete path is given by the sum of the costs of its steps. The cost of each step is given by the unlikelihood of the current state in the direction of motion of the current step. As the number of steps increases, in the limit we have a continuous path whose cost will be given by an integral as in (3.1).

3.3 Exit cost problem in coordination games

For $K, L \subset X$, we denote the set of all absolutely continuous paths of arbitrary duration through X from K to L by $\Phi(K, L)$. In the small noise double limit, the expected time until the process $\mathbf{X}^{N,\eta}$ exits from equilibrium e_i to another equilibrium is captured by the *cost of exit* given by

$$C(\{e_i\}, \cup_{j\neq i}\{e_j\}) = \min\{c(\phi) : \phi \in \Phi(\{e_i\}, \cup_{j\neq i}\{e_j\})\} \tag{3.2}$$

SS16 show that when N is sufficiently large, the exponential growth rate of the expected waiting time to leave the initial basin of attraction $\mathscr{B}^i$ as η vanishes is approximately $NC(\{e_i\}, \cup_{j\neq i}\{e_j\})$.

4. Preliminary Analysis

4.1 Definitions

4.1.1 Payoffs

We follow the notation introduced by SS16 for working with symmetric normal form games $A \in \mathbb{R}^{n\times n}$. We use superscripts to refer to rows of A and subscripts to refer to columns. Thus A^i is the ith row of A, A_j is the jth column of A, and A^i_j is the (i, j)th entry. These objects can be obtained by pre- and post-multiplying A by standard basis vectors:

$$A^i = e_i'A, \quad A_j = Ae_j, \quad A^i_j = e_i'Ae_j.$$

In a similar fashion, we use super- and subscripts of the form $i - j$ to denote certain differences obtained from A.

$$A^{i-j} = A^i - A^j = (e_i - e_j)'A, \quad A^{i-j}_{k-l} = A^i_k - A^i_l - A^j_k + A^j_l = (e_i - e_j)'A(e_k - e_l).$$

$$A^{2j-i-k}_{i-k} = A^{j-k}_{i-k} - A^{i-j}_{i-k}, \quad A^{2j-i-k}x = A^{j-k}x - A^{i-j}x$$
$$A^{2k-i-j}_{i-j} = A^{k-j}_{i-j} - A^{i-k}_{i-j}, \quad A^{2k-i-j}x = A^{k-j}x - A^{i-k}x$$

In this notation, the best response region for strategy i is described by

(4.1) $\qquad \mathcal{B}^i = \{x \in X : A^{i-l}x \geq 0 \text{ for all } l \in S\}.$

The set $\mathcal{B}^{ij} = \mathcal{B}^i \cap \mathcal{B}^j$ is the boundary between the best response regions for strategies i and j.

In the present notation, A is a *coordination game* (CG) if

(4.2) $\qquad A^i_i > A^j_i$ for all $i, j \in S$ with $j \neq i$,

so that each pure state is a Nash equilibrium of F. This implies that

(4.3) $\qquad A^{i-j}_{i-j} > 0$ for all $i, j \in S$.

We call A^{i-j}_{i-j} the (i, j)th *alignment* of A. This quantity, which corresponds to the denominator of the mixed equilibrium weights in the binary-choice game with strategies i and j, represents the strength of incentives to coordinate (or, if negative, to miscoordinate) in the restricted game with strategy set $\{i, j\}$.

Likewise, game A has the *marginal bandwagon property* (MBP) if

$$A^{i-j}_{i-k} > 0 \text{ for all } i, j, k \in S \text{ with } i \notin \{j, k\}. \tag{4.4}$$

4.1.2 Payoff Ranking Regions

Based on the form of the unlikelihood function for probit choice (see Lemma 2.3), we divide the best response region of strategy i, $\mathcal{B}^i$ as follows:

$$\begin{aligned}
\mathcal{B}^i_{jk} &= \{x \in \mathcal{B}^i : A^{j-k}x \geq 0\} \\
\mathcal{B}^i_{kj} &= \{x \in \mathcal{B}^i : A^{j-k}x \leq 0\} \\
\overline{\mathcal{B}}^i_{jk} &= \left\{x \in \mathcal{B}^i_{jk} : A^{2j-i-k}x \geq 0\right\} \\
\underline{\mathcal{B}}^i_{jk} &= \left\{x \in \mathcal{B}^i_{jk} : A^{2j-i-k}x \leq 0\right\} \\
\overline{\mathcal{B}}^i_{kj} &= \left\{x \in \mathcal{B}^i_{kj} : A^{2k-i-j}x \geq 0\right\} \\
\underline{\mathcal{B}}^i_{kj} &= \left\{x \in \mathcal{B}^i_{kj} : A^{2k-i-j}x \leq 0\right\}
\end{aligned} \tag{4.5}$$

In words, $\mathcal{B}^i_{jk}$ is the region in the best response of strategy i, where strategy j earns a payoff atleast as much as strategy k. In $\overline{\mathcal{B}}^i_{jk}$, strategy j also earns a payoff atleast as much as the average payoff of strategies i and k. Intuitively, in the interior of $\overline{\mathcal{B}}^i_{jk}$, strategy k earns a much lower payoff compared to strategies i and j. In $\underline{\mathcal{B}}^i_{jk}$, strategy j earns a payoff lower than the average payoff of strategies i and k. We can similarly interpret $\mathcal{B}^i_{kj}$, $\overline{\mathcal{B}}^i_{kj}$ and $\underline{\mathcal{B}}^i_{kj}$ by interchanging the roles of indices j and k.

We denote the lines as follows:

$$l^{ij} = \{se_i + (1-s)e_j : s \in \mathbb{R}\} \tag{4.6}$$

$$l^{ik} = \{se_i + (1-s)e_k : s \in \mathbb{R}\} \tag{4.7}$$

Let $\tilde{x}^{ij} = x^* + x^*_j(e_i - e_j)$ and $\tilde{x}^{ik} = x^* + x^*_k(e_i - e_k)$. Here $x^* = (x^*_i, x^*_j, x^*_k) \in \text{int}(X)$ is the unique completely mixed equilibrium.

4.1.3 Costs of direct paths

For $x \in \mathcal{B}^i$ with $x_k \leq x_k^*$, [2] we denote $W^j(x)$ to be the cost of a path from state x that moves through $\mathcal{B}^i$ in direction $e_j - e_i$ until reaching boundary $\mathcal{B}^{ij}$. Similarly, for $x \in \mathcal{B}^i$ with $x_j \leq x_j^*$, we denote $W^k(x)$ to be the cost of a path from state x that moves through $\mathcal{B}^i$ in direction $e_k - e_i$ until reaching boundary $\mathcal{B}^{ik}$. For $l = j, k$:

$$W^l(x) = \{c(\phi) \mid \phi : [0,1] \to \mathcal{B}^i \text{ such that } \phi(t) = x + t(y - x),\ \phi(0) = x, \quad \phi(1) = y \in \mathcal{B}^{il} \text{ and } y - x = d(e_l - e_i) \text{ for some } d > 0\} \tag{4.8}$$

We will compute these costs explicitly in Section 5. The solution to the *exit problem* will be expressed in terms of $W^j(x)$ and $W^k(x)$.

We summarize the definitions of important states on the affine hull of the state space aff(X) in Table 1 (also see Figure 1) which will be used extensively later in the main analysis in Section 6.

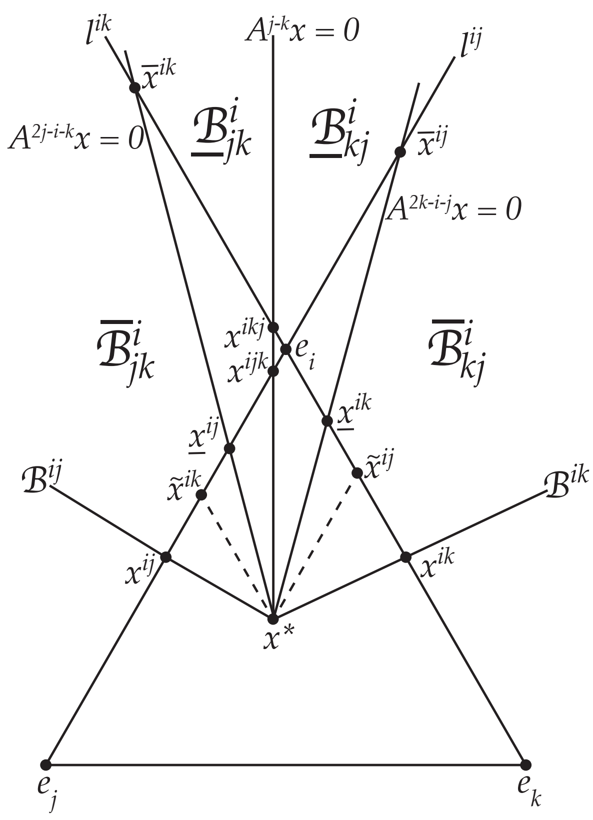


Figure 1: Division of best response region of strategy i

[2]The restriction $x_k \leq x_k^*$ is necessary for $W^j(x)$ to be defined. This is because, for $x \in \mathcal{B}^i$ with $x_k > x_k^*$, the path from x moving in $e_j - e_i$ will hit $\mathcal{B}^{ik}$ before reaching $\mathcal{B}^{ij}$ (see Figure 1). Similarly, the restriction $x_j \leq x_j^*$ is necessary for $W^k(x)$ to be defined.

Table 1: Description of Important States

Point	Description
x^*	unique interior mixed equlibrium of simple three strategy coordination game A
x^{ij}	unique mixed equilibrium of A with support $\{i, j\}$
x^{ik}	unique mixed equilibrium of A with support $\{i, k\}$
x^{ijk}	unique state on l^{ij} (see (4.6)) where payoffs to strategies j and k are equal
x^{ikj}	unique state on l^{ik} (see (4.7)) where payoffs to strategies j and k are equal
$\underline{x}^{ij}$	unique state on l^{ij} where the average payoff of strategies i and k is equal to the payoff of j
$\overline{x}^{ij}$	unique state on l^{ij} where the average payoff of strategies i and j is equal to the payoff of k
$\underline{x}^{ik}$	unique state on l^{ik} where the average payoff of strategies i and j is equal to the payoff of k
$\overline{x}^{ik}$	unique state on l^{ik} where the average payoff of strategies i and k is equal to the payoff of j

4.2 Outline of analysis and main results

The *exit problem* (3.2), has nonsmooth running cost, and is multidimensional in games with more than two strategies. Nevertheless, this problem can be solved explicitly by using verification theorem from optimal control theory. This involves first guessing a continuous, piecewise cubic value function and then checking a Hamilton Jacobi Bellman (HJB) equation (5.2) in certain basic directions of motion. In our case, we have six possible basic directions of motion of the form $e_b - e_a$, where $a, b \in \{1, 2, 3\}$ and $b \neq a$.

In Section 5.1, we provide the statement of the verification theorem which gives a sufficient condition for a function to be a value function for a specific class of optimization problems which includes the exit problem. We then solve the exit problem in the following sequence of steps. Using the unlikelihood function for the probit choice, we compute the costs of certain direct paths using (3.1) in Section 5.2. These costs will be used to compute $W^j(x)$ and $W^k(x)$ (see (4.8) for definitions). In Section 5.3, we show that for states x in the interior of $\mathcal{B}^i_{jk}$, the value function generated by moving in the direction $e_j - e_i$ until reaching $\mathcal{B}^{ij}$ satisfies the HJB equation (see Lemma 5.5). We then analyze the behavior of the function $W^k(x) - W^j(x)$ on the lines l^{ij} and l^{ik}. Using a homogeneity argument we show that there exists a unique state $\hat{x}^i \in \mathcal{B}^i \cap \mathrm{bd}(X)$ such that $W^k(\hat{x}^i) = W^j(\hat{x}^i)$ (see Proposition 5.8). We conclude Section 5 by showing that $\hat{x}^i$ is in one of the regions (4.5) of $\mathcal{B}^i$ depending on the entries of the matrix A as summarized in Table 2. In Section 6, we show that the optimal solution to the exit problem can be broadly divided into the following cases:

Table 2: Description of Different Cases

Case	Condition on the entries of A
$\hat{x}^i \in \underline{\mathcal{B}}^i_{jk} \cup \underline{\mathcal{B}}^i_{kj}$	$A^{i-k}_{i-k} \leq 8A^{i-j}_{i-j}$ and $A^{i-j}_{i-j} \leq 8A^{i-k}_{i-k}$
$\hat{x}^i \in (\overline{\mathcal{B}}^i_{jk})^\circ$	$A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$
$\hat{x}^i \in (\overline{\mathcal{B}}^i_{kj})^\circ$	$A^{i-j}_{i-j} > 8A^{i-k}_{i-k}$

4.2.1 Standard case

When (i) $A^{i-k}_{i-k} \leq 8A^{i-j}_{i-j}$ and $A^{i-j}_{i-j} \leq 8A^{i-k}_{i-k}$ or (ii) $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$ and $DW^k(\hat{x}^i)(e_i - e_j) \geq 0$ or (iii) $A^{i-j}_{i-j} > 8A^{i-k}_{i-k}$ and $DW^j(\hat{x}^i)(e_i - e_k) \geq 0$, we show that the optimal exit paths divide the best response region $\mathcal{B}^i$ into two regions; in one the optimal control is $e_j - e_i$, and the exit path leads to $\mathcal{B}^{ij}$; in the other the optimal control is $e_k - e_i$, and the exit path leads to $\mathcal{B}^{ik}$ (see Figure 2). The solution in this case is qualitatively similar to the case where agents follow the logit choice (see Proposition 12 of SS16). We call this the *standard case*.

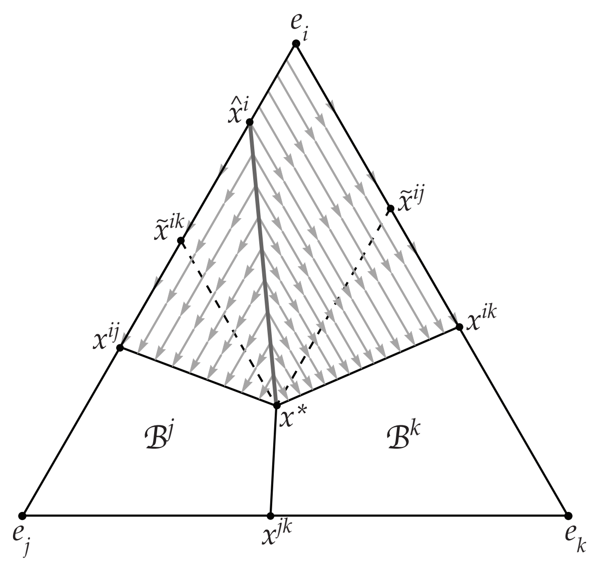


Figure 2: Optimal exit paths in standard cases from $\mathcal{B}^i$ when $\hat{x}^i$ is on the face e_ie_j.

4.2.2 Retreating case

When $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$ and $DW^k(\hat{x}^i)(e_i - e_j) < 0$, the optimal exit paths divide the best response region $\mathcal{B}^i$ into three regions; in one the optimal control is $e_j - e_i$, and exit path leads to $\mathcal{B}^{ij}$; in one the optimal control is $e_k - e_i$, and exit path leads to $\mathcal{B}^{ik}$ and between these two regions there is a third region in which the optimal control is $e_i - e_j$. Similarly,

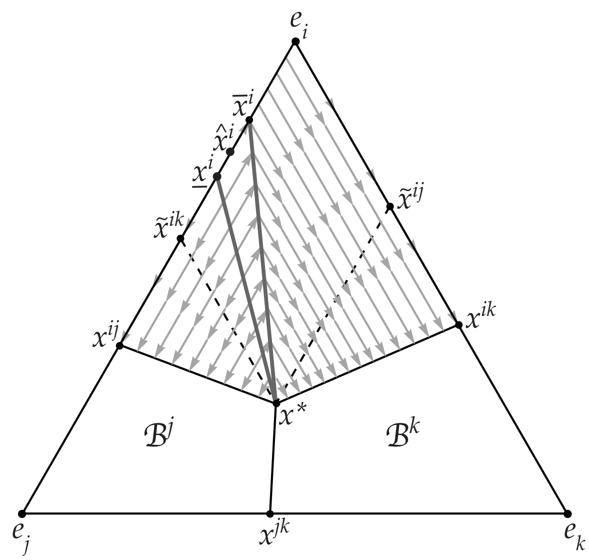


Figure 3: Optimal exit paths in retreating cases with non-binding state constraints from $\mathcal{B}^i$ when $\hat{x}^i$ is in $\overline{\mathcal{B}}^i_{jk}$ and on the face e_ie_j.

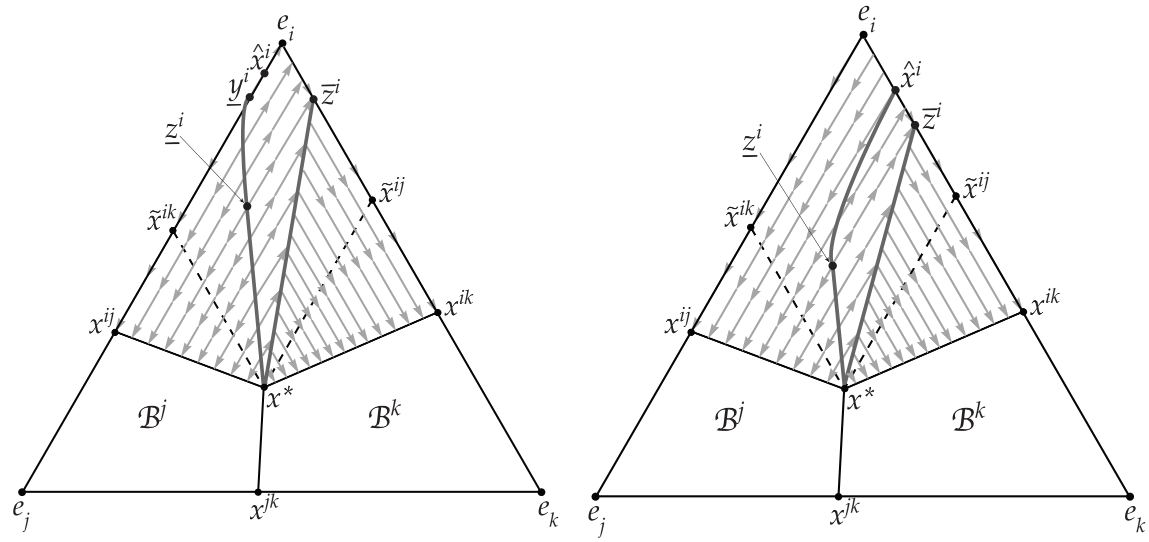
when $A^{i-j}_{i-j} > 8A^{i-k}_{i-k}$ and $DW^j(\hat{x}^i)(e_i - e_k) < 0$, the optimal exit paths divide the best response region $\mathcal{B}^i$ into three regions; in one the optimal control is $e_j - e_i$, and exit path leads to $\mathcal{B}^{ij}$; in one the optimal control is $e_k - e_i$, and exit path leads to $\mathcal{B}^{ik}$ and between these two regions there is a third region in which the optimal control is $e_i - e_k$. We call these *retreating cases* as there is a region in which the optimal exit path heads back in the direction of the initial equilibrium e_i. Our main analysis focuses only on the cases with $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$ and $DW^k(\hat{x}^i)(e_i - e_j) < 0$. This is without loss of generality because interchanging the roles of indices j and k covers the cases with $A^{i-j}_{i-j} > 8A^{i-k}_{i-k}$ and $DW^j(\hat{x}^i)(e_i - e_k) < 0$.

The retreating case can be further divided into two subcases: optimal exit paths with non-binding state constraints as shown in Figure 3 (see Theorem 6.7), optimal exit paths with binding state constraints as shown in Figure 4 (see Theorem 6.8).

5. Main Analysis

5.1 Verification Theorem

Let TX denote the set of tangent vectors from states in the relative interior X° of X. Let the set $\Omega \subset X$ be closed relative to X and have piecewise smooth boundary. Let the function $L : X \times TX \to \mathbb{R}_+$ be a semi-linear Lipschitz continuous and $\mathcal{Z} \subset TX$ be compact and convex. The control problem and its value function $V^* : X \to \mathbb{R}_+$ are defined as follows:



(i) Optimal exit paths in retreating cases from $\mathcal{B}^i$ when $\hat{x}^i$ is in $\overline{\mathcal{B}}^i_{jk}$ and on the face e_ie_j

(ii) Optimal exit paths in retreating cases from $\mathcal{B}^i$ when $\hat{x}^i$ is in $\overline{\mathcal{B}}^i_{jk}$ and on the face e_ie_k

Figure 4: Binding state constraints

$$V^*(x) = \min \int_0^T L(\phi_t, \nu_t)dt \tag{5.1}$$

over $\qquad T \in [0, \infty), \nu : [0, T] \to \mathcal{Z}$ measurable
subject to $\phi : [0, T] \to X$ absolutely continuous, $\phi_0 = x, \phi_T \in \Omega, \dot{\phi}_t = \nu_t$ for almost every $t \in [0, T]$.

The results proved in Sandholm et al. (2018) provide a sufficient condition for a function $V : X \to \mathbb{R}_+$ to be the value function of the above problem. The key requirement is that the *Hamilton-Jacobi-Bellman(HJB) equation*

$$\min_{u \in \mathcal{Z}}(L(x, u) + DV(x)u) = 0 \tag{5.2}$$

hold at almost every $x \in X^\circ$. Adapting Theorems 3.3 and 3.4 from Sandholm et al. (2018) to our specific problem gives the following verification theorem.

Theorem 5.1. (Verification theorem (Sandholm et al. (2018))). *Let $V : X \to \mathbb{R}_+$ be a continuous function that is continuously differentiable a.e. on X°. Suppose that*

(i) For every $x \in X$, there is a time $T \in [0, \infty)$ and a measurable function $\nu : [0, T] \to TX$ such that the corresponding controlled trajectory $\phi : [0, T] \to X$ with $\phi_0 = x$ satisfies $\phi_T \in \Omega$ and $\int_0^T L(\phi_t, \nu_t)dt = V(x)$;

(ii) The HJB equation (5.2) holds a.e. on X°.

(iii) The boundary condition $V(x) = 0$ holds at all $x \in \Omega$.

Then $V = V^*$.

We will use the above theorem in Section 6 to verify the solution to the exit problem.

5.2 Costs of direct paths

Let $\gamma(x, y)$ denote the cost of the direct path from x to y :

(5.3) $\qquad \gamma(x, y) = c(\phi),\;$ where $\phi : [0, 1] \to X$ is defined by $\phi_t = (1 - t)x + ty.$

Lemma 5.2. *Let $x \in \mathcal{B}^i_{jk}$ and suppose that*

(5.4) $\qquad y = x + d(e_j - e_i) \in \mathcal{B}^{ij}$ *for some* $d > 0$

and that $A^{i-j}_{i-j} \neq 0$. Then $\gamma(x, y) = \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}}$

Proof. Since $y \in \mathcal{B}^{ij}$, we have $A^{i-j}y = 0$ and therefore $d = \frac{A^{i-j}x}{A^{i-j}_{i-j}}$ from (5.4). Let $\pi = F(\phi_t) = A\phi_t$, where $\phi_t = x + t(y - x)$. For $l = i, j$ and k we have

$$F_l(\phi_t) = A^l(x + t(y - x)) = (1 - t)A^l x + tA^l y$$

Clearly, $\dot{\phi}_t = y - x = d(e_j - e_i)$ and therefore $[\dot{\phi}_t]'_+ = de'_j$. $x, y \in \mathcal{B}^i_{jk}$ implies that $\phi_t \in \mathcal{B}^i_{jk}$. Therefore, from Lemma 2.3, we have

$$e'_j\Upsilon(F(\phi_t)) = \frac{1}{4}(F_i(\phi_t) - F_j(\phi_t))^2$$

We now compute as follows:

$$\begin{aligned}
\gamma(x, y) &= \int_0^1 [\dot{\phi}_t]'_+\Upsilon(F(\phi_t))dt \\
&= \int_0^1 de'_j\Upsilon(F(\phi_t))dt \\
&= \frac{d}{4}\int_0^1 (F_i(\phi_t) - F_j(\phi_t))^2 dt \\
&= \frac{d}{4}\int_0^1 ((1 - t)A^{i-j}x + tA^{i-j}y)^2 dt \\
&= \frac{d}{4}\int_0^1 (1 - t)^2(A^{i-j}x)^2 dt \qquad\qquad \text{(since } A^{i-j}y = 0\text{)} \\
&= \frac{d}{4} \times \frac{1}{3}(A^{i-j}x)^2
\end{aligned}$$

$$= \frac{d}{12}(A^{i-j}x)^2$$

$$= \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}} \qquad \left(\text{since } d = \frac{A^{i-j}x}{A^{i-j}_{i-j}}\right)$$

□

Lemma 5.3. *Let $x \in \mathcal{B}^i_{jk}$ with $x_j < x^*_j$ and suppose that*

(5.5) $\quad y = x + d(e_k - e_i) \in \mathcal{B}^{ik}$ *for some* $d > 0$

and that $A^{i-k}_{i-k} \neq 0$.

(i) If $x \in \underline{\mathcal{B}}^i_{jk}$, then $\gamma(x, y) = \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}}$.

(ii) If $x \in \overline{\mathcal{B}}^i_{jk}$, then $A^{2j-i-k}_{i-k} > 0$ and $\gamma(x, y) = \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} + \frac{1}{36}\frac{(A^{2j-i-k}x)^3}{A^{2j-i-k}_{i-k}}$

Proof. Since $y \in \mathcal{B}^{ik}$, we have $A^{i-k}y = 0$ and $d = \frac{A^{i-k}x}{A^{i-k}_{i-k}}$. Following similar steps as in the proof of Lemma 5.2, we let $\pi = F(\phi_t) = A\phi_t$, where $\phi_t = x + t(y - x)$. Clearly, $\dot{\phi}_t = y - x = d(e_k - e_i)$ and therefore $[\dot{\phi}_t]'_+ = de'_k$.

For $x \in \underline{\mathcal{B}}^i_{jk}$, since $y \in \mathcal{B}^{ik}$, we have $\phi_t \notin \overline{\mathcal{B}}^i_{jk}$ (see Figure 1) for any $t \in [0, 1]$. Therefore from (4.5) and Lemma 2.3, we have $e'_k\Upsilon(F(\phi_t)) = \frac{1}{4}(F_i(\phi_t) - F_k(\phi_t))^2$. Following the computations as in Lemma 5.2 by replacing the index j with k, we get $\gamma(x, y) = \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}}$.

The computations for $x \in \overline{\mathcal{B}}^i_{jk}$ are in Appendix B. □

For $x \in \overline{\mathcal{B}}^i_{jk}$, the corresponding payoff vector $\pi = Ax$ is such that the payoff to strategy j is greater than the average payoff of strategies i and k. Since, we are in the best response region of strategy i, the above statement means that for states in $\overline{\mathcal{B}}^i_{jk}$, strategy k does badly compared to strategies i and j. This makes the movement along the direction $e_k - e_i$ unattractive and hence the additional term $\frac{1}{36}\frac{(A^{2j-i-k}x)^3}{A^{2j-i-k}_{i-k}}$ in the cost function for states $x \in \overline{\mathcal{B}}^i_{jk}$.

Remark 5.4. When the line $\frac{\pi_k+\pi_i}{2} = \pi_j$ is below the line joining the states x^* and $\tilde{x}^{ik}$ then there does not exist any $x \in \overline{\mathcal{B}}^i_{jk}$ with $x_j < x^*_j$. Lemma 5.3(ii) does not apply in such cases. It holds only when there exists some $x \in \overline{\mathcal{B}}^i_{jk}$ with $x \in \mathcal{B}^i$ and $x_j < x^*_j$, in which case, we have $A^{2j-i-k}_{i-k} > 0$. Similarly, for $x \in \mathcal{B}^i$ with $x_k < x^*_k$, if there exists some $x \in \overline{\mathcal{B}}^i_{kj}$ then $A^{2k-i-j}_{i-j} > 0$.

For $x \in \mathcal{B}^i$ with $x_j < x^*_j$, from Lemmas 5.2 and 5.3, we have

$$W^k(x) = \begin{cases} \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} & \text{if } x \notin \overline{\mathcal{B}}^i_{jk} \\ \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} + \frac{1}{36}\frac{(A^{2j-i-k}x)^3}{A^{2j-i-k}_{i-k}} & \text{if } x \in \overline{\mathcal{B}}^i_{jk} \end{cases} \tag{5.6}$$

By interchanging indices j and k, for $x \in \mathcal{B}^i$ with $x_k < x^*_k$, we have

$$W^j(x) = \begin{cases} \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}} & \text{if } x \notin \overline{\mathcal{B}}^i_{kj} \\ \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}} + \frac{1}{36}\frac{(A^{2k-i-j}x)^3}{A^{2k-i-j}_{i-j}} & \text{if } x \in \overline{\mathcal{B}}^i_{kj} \end{cases} \tag{5.7}$$

5.3 The Value function for the exit cost problem

To solve the *exit problem* via dynamic programming, we first determine the form of the value function at states near the target set. We therefore consider the cost of reaching the set $\mathcal{B}^j$ from nearby states in $\mathcal{B}^i$. It is natural to guess that there is a region $R^{ij} \subseteq \mathcal{B}^i$ whose boundary contains $\mathcal{B}^{ij}$ in which motion in direction $e_j - e_i$ leads to $\mathcal{B}^{ij}$, and in fact defines the optimal feedback control. By Lemma 5.2, this choice of control generates the candidate value function

$$V(x) = \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}} \tag{5.8}$$

in region R^{ij}.

Lemma 5.5. *Suppose that the function V is defined by equation* (5.8) *on $R^{ij} \subseteq \mathcal{B}^i$ as specified above. Then the HJB equation* (5.2) *for V is satisfied for all $x \in (\mathcal{B}^i_{jk})^\circ$.*

Proof. For $x \in \overline{\mathcal{B}}^i_{jk}$, the running cost function $L(x,u) = (\Upsilon(F(x)))'[u]_+$ for the probit choice rule is given by

$$L(x,u) = \left[0 \quad \frac{1}{4}(A^{i-j}x)^2 \quad \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2\right)\right][u]_+ \tag{5.9}$$

Similarly, for $x \in \underline{\mathcal{B}}^i_{jk}$, the running cost function is given by

$$L(x,u) = \left[0 \quad \frac{1}{4}(A^{i-j}x)^2 \quad \frac{1}{4}(A^{i-k}x)^2\right][u]_+ \tag{5.10}$$

Differentiating (5.8), we get $DV(x) = \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}}A^{i-j}$. Let $H(x,u) = L(x,u) + DV(x)u$. In order to verify the (HJB) equation (5.2), it is sufficient to show that the function $H(x,u)$ attains a value of 0 when $u = e_j - e_i$ and for the remaining five basic directions $e_i - e_j$, $e_j - e_k$, $e_k - e_j$, $e_i - e_k$ and $e_k - e_i$, $H(x, u) \geq 0$. We verify this for states $x \in (\overline{\mathcal{B}}^i_{jk})^\circ$ and $x \in (\underline{\mathcal{B}}^i_{jk})^\circ$ in Appendix C. □

Lemma 5.6. *For a simple three-strategy coordination game A, the line $\pi_j = \pi_k$ is between the lines $x^*\tilde{x}^{ik}$ and $x^*\tilde{x}^{ij}$.*

Proof. We first show that the line $\pi_j = \pi_k$ is above the line $x^*\tilde{x}^{ik}$. It suffices to show that $A^j\tilde{x}^{ik} > A^k\tilde{x}^{ik}$. We compute as follows

$$\begin{aligned}
A^{j-k}\tilde{x}^{ik} &= A^{j-k}(x^* + x^*_k(e_i - e_k)) \\
&= A^{j-k}x^* + x^*_k A^{j-k}(e_i - e_k) \\
&= x^*_k A^{j-k}_{i-k} \\
&= 0 \qquad\qquad \text{(by MBP (4.4))}
\end{aligned}$$

Similarly, we can show that the line $\pi_k = \pi_j$ is above the line $x^*\tilde{x}^{ij}$ (see Figure 1). □

Using Lemmas 5.5 and 5.6, we will show in Section 6 that the optimal exit path for a state $x \in \mathcal{B}^i$ with $x_j \geq x^*_j$ is moving in direction $e_j - e_i$ until reaching $\mathcal{B}^{ij}$.

Theorem 5.7. *There is a unique state $z^{ij} \in l^{ij}$ with $z^{ij}_i > \tilde{x}^{ik}_i$ such that $W^k(z^{ij}) = W^j(z^{ij})$ and a unique state $z^{ik} \in l^{ik}$ with $z^{ik}_i > \tilde{x}^{ij}_i$ such that $W^k(z^{ik}) = W^j(z^{ik})$.*

Proof. See Appendix D. □

Proposition 5.8. *There is a unique state $\hat{x}^i \in \mathcal{B}^i \cap bd(X)$ such that $\hat{x}^i_j < x^*_j$, $\hat{x}^i_k < x^*_k$ and $W^k(\hat{x}^i) = W^j(\hat{x}^i)$.*

Proof. From Theorem 5.7, we have a unique state $z^{ij} \in l^{ij}$ with $z^{ij}_i > \tilde{x}^{ik}_i$ i.e., $z^{ij}_j = 1 - z^{ij}_i < 1 - \tilde{x}^{ik}_i = x^*_j$, such that $W^k(z^{ij}) = W^j(z^{ij})$. Similarly, there is a unique state $z^{ik} \in l^{ik}$ with $z^{ik}_k < x^*_k$ such that $W^k(z^{ik}) = W^j(z^{ik})$. To complete the proof we establish homogeneity of degree 3 of $W^k(x)$ and $W^j(x)$ in the displacement vector $z = x - x^*$ of x from x^*. For $t \in \mathbb{R}_+$, $x^* + z \in \overline{\mathcal{B}}^i_{jk}$, we have $x^* + tz \in \overline{\mathcal{B}}^i_{jk}$. For $x^* + tz \in \overline{\mathcal{B}}^i_{jk}$, from Lemma 5.3, we have

$$\begin{aligned}
W^k(x^* + tz) &= \frac{1}{12}\frac{(A^{i-k}(x^* + tz))^3}{A^{i-k}_{i-k}} + \frac{1}{36}\frac{(A^{2j-i-k}(x^* + tz))^3}{A^{2j-i-k}_{i-k}} \\
&= \frac{1}{12}\frac{(A^{i-k}(tz))^3}{A^{i-k}_{i-k}} + \frac{1}{36}\frac{(A^{2j-i-k}(tz))^3}{A^{2j-i-k}_{i-k}}
\end{aligned}$$

$$= \frac{t^3}{12}\frac{(A^{i-k}z)^3}{A^{i-k}_{i-k}} + \frac{t^3}{36}\frac{(A^{2j-i-k}z)^3}{A^{2j-i-k}_{i-k}}$$
$$= \frac{t^3}{12}\frac{(A^{i-k}(x^*+z))^3}{A^{i-k}_{i-k}} + \frac{t^3}{36}\frac{(A^{2j-i-k}(x^*+z))^3}{A^{2j-i-k}_{i-k}}$$
$$= t^3 W^k(x^*+z) \tag{5.11}$$

Following similar computations as above, we can show that for $x^* + tz \in \mathcal{B}^i \setminus \overline{\mathcal{B}}^i_{jk}$, $W^k(x^* + tz) = t^3 W^k(x^* + z)$ for all $t \in \mathbb{R}_+$. It follows that $W^k(x^* + tz) = t^3 W^k(x^* + z)$ for $x^* + z \in \mathcal{B}^i$, for all $t \in \mathbb{R}_+$. Similarly, $W^j(x^* + tz) = t^3 W^j(x^* + z)$ for $x^* + z \in \mathcal{B}^i$, for all $t \in \mathbb{R}_+$. We therefore have,

$$W^k(x^* + tz) - W^j(x^* + tz) = t^3(W^k(x^* + z) - W^j(x^* + z))$$

for $x^* + z \in \mathcal{B}^i$, for all $t \in \mathbb{R}_+$.

Thus if $W^k(x^* + z) = W^j(x^* + z)$, then $W^k(x^* + tz) = W^j(x^* + tz)$ for all $t \in \mathbb{R}_+$. It therefore follows that z^{ij}, z^{ik} and x^* are collinear. If z^{ij} and z^{ik} are both e_i, we set $\hat{x}^i = e_i$. If not, exactly one of z^{ij} and z^{ik} is in X, and that is our $\hat{x}^i$. □

Lemma 5.9. *$\hat{x}^i$ lies in $(\overline{\mathcal{B}}^i_{jk})^\circ$ iff $A^{i-k}_{i-k} \geq 8A^{i-j}_{i-j}$.*

Proof. For $\hat{x}^i$ to lie in $\overline{\mathcal{B}}^i_{jk}$, the unique root $s_0 > \tilde{x}^{ik}_i$ which solves the equation $g(s) = 0$ (see equation (D.1)) is such that $s_0 < \underline{x}^{ij}_i$. Since $g(\tilde{x}^{ik}) > 0$ (this follows from (D.9)), this implies that we should have $g(\underline{x}^{ij}_i) < 0$.

$$\begin{aligned} g(\underline{x}^{ij}_i) &= \frac{(A^{i-k}\underline{x}^{ij})^3}{A^{i-k}_{i-k}} - \frac{(A^{i-j}\underline{x}^{ij})^3}{A^{i-j}_{i-j}} \\ &= \frac{(2A^{i-j}\underline{x}^{ij})^3}{A^{i-k}_{i-k}} - \frac{(A^{i-j}\underline{x}^{ij})^3}{A^{i-j}_{i-j}} && \text{(since, by definition } A^{j-k}\underline{x}^{ij} = A^{i-j}\underline{x}^{ij}) \\ &= (A^{i-j}\underline{x}^{ij})^3\left(\frac{8}{A^{i-k}_{i-k}} - \frac{1}{A^{i-j}_{i-j}}\right) \end{aligned}$$

From the above set of equations, it follows that $g(\underline{x}^{ij}_i) < 0$ implies $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$. □

Interchanging the roles of indices j and k in Lemma 5.9, we have the following:

Corollary 5.10. *$\hat{x}^i$ lies in $(\overline{\mathcal{B}}^i_{kj})^\circ$ iff $A^{i-j}_{i-j} > 8A^{i-k}_{i-k}$.*

From Lemma 5.9 and Corollary 5.10, we have the following:

Corollary 5.11. *$\hat{x}^i$ lies in $\underline{\mathcal{B}}^i_{jk} \cup \underline{\mathcal{B}}^i_{kj}$ iff $A^{i-k}_{i-k} \leq 8A^{i-j}_{i-j}$ and $A^{i-j}_{i-j} \leq 8A^{i-k}_{i-k}$.*

We are now ready to state and prove the solution to the exit problem under probit choice in the next section. Under logit choice for a simple three strategy coordination game, the optimal exit path divides the initial basin of attraction into two regions as shown in Figure 2. However under probit choice the solution to the exit problem has many possible forms, depending on the relative magnitudes of coordination incentives between i and j (as measured by A^{i-j}_{i-j}) and between i and k (as measured by A^{i-k}_{i-k}). We consider this in detail in Section 6.

6. Main Results

As pointed in Section 5, the key requirement for a function V to be a value function is that the HJB equation (5.2) holds at almost every $x \in X^\circ$. Verifying this at almost every x in the interior of X can be quite cumbersome. By extending the state space X to the affine hull $\mathrm{aff}(X) = \{x \in \mathbb{R}^n : \sum_i x_i = 1\}$,[3] we can restrict the number of states at which we verify the HJB equation (5.2) significantly as we show next. Let

$$H(x, u) = L(x, u) + DV(x)u \tag{6.1}$$

In (5.11), we established that $W^k(x^* + tz)$ is homogeneous of degree 3 in t for $t \geq 0$. This implies that its derivative $DW^k(x^* + tz)$ is homogeneous of degree 2 in t for $t \geq 0$. Similarly, $DW^j(x^* + tz)$ is homogeneous of degree 2 in t for $t \geq 0$. The value function V for the exit problem which we verify will either be $W^k(x)$ or $W^j(x)$ depending on x and so it follows that $DV(x^* + tz)$ is homogeneous of degree 2 in t for $t \geq 0$. It is straightforward to verify from (5.9) and (5.10) that the running cost $L(x, u)$ satisfies $L(x^* + tz, u) = t^2 L(x^* + z, u)$ for $z = x - x^*$ and for all $t \in \mathbb{R}_+$. Since $H(x^* + tz, u) = L(x^* + tz, u) + DV(x^* + tz, u)$, it follows that

$$H(x^* + tz, u) = t^2 H(x^* + z, u) \tag{6.2}$$

For $z = x - x^*$ with x on l^{ij} it is clear from (6.2) that $H(x^* + z, u) \geq 0$ implies $H(x^* + tz, u) \geq 0$. Therefore it is sufficient to verify the HJB equation (5.2) on the line l^{ij}.

[3] In the extended state space, $\mathcal{B}^i = \{x \in \mathrm{aff}(X) : A^{i-l}x \geq 0 \text{ for all } l \in S\}$ and the payoff ranking regions in Section 4.1.2 will be defined with respect to this.

6.1 Standard Case

Theorem 6.1. *Let A be a simple three-strategy coordination game. Suppose $\hat{x}^i$ is in $\underline{\mathcal{B}}^i_{jk} \cup \underline{\mathcal{B}}^i_{kj}$ i.e., $A^{i-j}_{i-j} \leq 8A^{i-k}_{i-k}$ and $A^{i-k}_{i-k} \leq 8A^{i-j}_{i-j}$. Let $v^i = x^* \times (\hat{x}^i - x^*)$. In this case, the value function $W^* : \mathcal{B}^i \to \mathbb{R}_+$ for the exit cost problem with target set $\mathcal{B}^j \cup \mathcal{B}^k$ is given by the continuous function*

$$W^*(x) = \begin{cases} W^k(x) & \text{if } (v^i)'(x) \leq 0, \\ W^j(x) & \text{if } (v^i)'(x) > 0. \end{cases} \tag{6.3}$$

The optimal feedback controls in this case are (see Figure 5)

$$\nu^*(x) = \begin{cases} = e_k - e_i & \text{if } (v^i)'(x) < 0 \\ \in \{e_k - e_i, e_j - e_i\} & \text{if } (v^i)'(x) = 0, \\ = e_j - e_i & \text{if } (v^i)'(x) > 0. \end{cases} \tag{6.4}$$

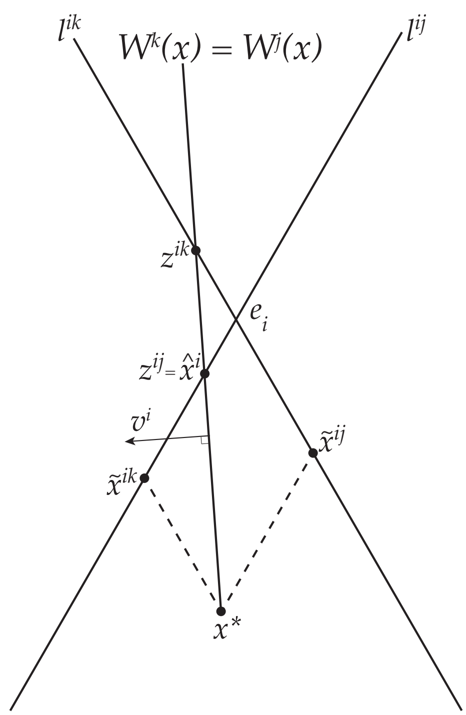


Figure 5: Standard case

Proof. We prove the theorem by assuming that $\hat{x}^i$ is in $\underline{\mathcal{B}}^i_{jk}$. The proof when $\hat{x}^i$ is in $\underline{\mathcal{B}}^i_{kj}$ is identical. First suppose that $\hat{x}^i$ is on $l^{ij} \cap X$. We established in Lemma 5.5 that for $x \in (\mathcal{B}^i_{jk})^\circ$, moving in the direction $e_j - e_i$ until reaching $\mathcal{B}^{ij}$ is consistent with the HJB equation. Relabeling j and k in Lemma 5.5 we have for $x \in (\mathcal{B}^i_{kj})^\circ$, moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ is consistent with the HJB equation. It remains to be shown that for $x \in (\mathcal{B}^i_{jk})^\circ$ with $W^k(x) < W^j(x)$ moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ is consistent with the HJB equation (5.2). Homogeneity argument (6.2) implies that it is sufficient to verify this at the states x on l^{ij} with $\hat{x}^i_i < x_i < x^{ijk}_i$. We prove this in Appendix E.

Next suppose that $\hat{x}^i$ is on $l^{ik} \cap X$. Following the previous arguments, it remains to be shown that for states x on l^{ik} with $x^{ikj}_i < x_i < \hat{x}^i_i$ moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ is consistent with the HJB equation. Another application of homogeneity argument (6.2) shows that this is equivalent to proving that for states x on l^{ij} with $z^{ij}_i < x_i < x^{ijk}_i$ moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ is consistent with the HJB equation (5.2). This follows from computations in Appendix E by replacing $\hat{x}^i$ with z^{ij}.

We now apply the Verification Theorem 5.1. The value function W^* in (6.3) is constructed from feedback controls (6.4) that generate feasible solutions to the exit problem as required by condition (i) of Theorem 5.1. The continuity of W^* follows from Proposition 5.8. W^* is clearly C^1 off the set $\{x \in X : (v^i)'x = 0\}$, and the arguments above imply that the HJB equation holds *a.e.* away from this set. Thus condition (ii) of Theorem 5.1 is satisfied and the proof is complete. The optimal solution to the exit problem in this case divides the state space into two regions as shown in Figure 2. □

Theorem 6.2. *Let A be a simple three-strategy coordination game. Suppose $\hat{x}^i \in (\overline{\mathcal{B}}^i_{jk})^\circ$ i.e., $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$. If $DW^k(\hat{x}^i)(e_i - e_j) \geq 0$, the value function $W^* : \mathcal{B}^i \to \mathbb{R}_+$ for the exit cost problem with target set $\mathcal{B}^j \cup \mathcal{B}^k$ is given by* (6.3) *and the optimal feedback controls by* (6.4)*.*

Proof. First suppose that $\hat{x}^i$ is on $l^{ij} \cap X$. Following the same arguments as in the proof of Theorem 6.1, we need to show that for x on l^{ij} with $\hat{x}^i_i \leq x_i < x^{ijk}_i$ moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ is consistent with the HJB equation. Consider the states x on l^{ij} with $\underline{x}^{ij}_i < x_i < x^{ijk}_i$. By definition, for such states $x \in \underline{\mathcal{B}}^i_{jk}$. From Lemma 5.3, we have $W^k(x) = \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}}$. Following the computations in Appendix E, it follows that the HJB equation is satisfied for such x.

We now consider the states x on l^{ij} with $\tilde{x}^{ik}_i < \hat{x}^i_i \leq x_i < \underline{x}^{ij}_i$. By definition, for such states, we have $x \in \overline{\mathcal{B}}^i_{jk}$. From Lemma 5.3,

$$W^k(x) = \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} + \frac{1}{36}\frac{(A^{2j-i-k}x)^3}{A^{2j-i-k}_{i-k}}$$

We compute (5.2) for such states x and show that $H(x, e_k - e_i) = 0$, $H(x, e_i - e_k) \geq 0$, $H(x, e_j - e_k) \geq 0$, $H(x, e_j - e_i) \geq 0$ and $H(x, e_k - e_j) = H(x, e_i - e_j) = DW^k(x)(e_i - e_j)$ in Appendix F. For states x with $\hat{x}^i_i \leq x_i < \underline{x}^{ij}_i$, the function $DW^k(x)(e_i - e_j)$ attains its minimum at $x = \hat{x}^i$. To prove this, define a function $c(s)$ as follows

$$(6.5) \qquad c(s) = DW^k(x(s))(e_i - e_j) \text{ for } x(s) = se_i + (1-s)e_j,\ \tilde{x}_i^{ik} \le s < \underline{x}_i^{ij}$$

Note that

$$DW^k(x) = \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}$$

Differentiating $c(s)$ using the above equation, we get

$$c'(s) = \frac{1}{2}\frac{A^{i-k}x}{A^{i-k}_{i-k}}(A^{i-k}_{i-j})^2 + \frac{1}{6}\frac{A^{2j-i-k}x}{A^{2j-i-k}_{i-k}}(A^{2j-i-k}_{i-j})^2$$

$x \in (\overline{\mathscr{B}}^i_{jk})^\circ$ implies $A^{2j-i-k}x > 0$ and from Lemma 5.3, we have $A^{2j-i-k}_{i-k} > 0$. Since $A^{i-k}_{i-k} > 0$ from CG (4.3), we conclude that

$$(6.6) \qquad c'(s) > 0 \text{ for } \tilde{x}_i^{ik} < s < \underline{x}_i^{ij}$$

Since $\tilde{x}_i^{ik} < \hat{x}_i^i \le \underline{x}_i^{ij}$, from (6.5) and (6.6) we conclude that the function $DW^k(x)(e_i - e_j)$ evaluated at the states x with $\hat{x}_i^i \le x < \underline{x}_i^{ij}$ attains its minimum at $x = \hat{x}^i$. Therefore, if $DW^k(\hat{x}^i)(e_i - e_j) \ge 0$, the HJB equation is satisfied by the function W^* for x with $\hat{x}_i^i \le x < \underline{x}_i^{ij}$.

Next suppose that $\hat{x}^i$ is on $l^{ik} \cap X$. Following the previous arguments, it remains to be shown that for states x on l^{ik} with $x_i^{ikj} < x_i \le \hat{x}_i^i$ moving in the direction $e_k - e_i$ until reaching $\mathscr{B}^{ik}$ is consistent with the HJB equation. An application of homogeneity argument (6.2) shows that this is equivalent to showing that for states x on l^{ij} with $z_i^{ij} \le x_i < x_i^{ijk}$ moving in the direction $e_k - e_i$ until reaching $\mathscr{B}^{ik}$ is consistent with the HJB equation. This follows from the preceding arguments by replacing $\hat{x}^i$ with z^{ij}.

We now apply the verification theorem. The value function W^* in (6.3) is constructed from feedback controls (6.4) that generate feasible solutions to the exit problem, as required by condition (i) of Theorem 5.1. The continuity of W^* follows from Proposition 5.8. W^* is clearly C^1 off the set $\{x \in \mathrm{aff}(X) : (v^i)'x = 0\}$, and the arguments above imply that the HJB equation holds away *a.e.* from this set. Thus condition (ii) of Theorem 5.1 is satisfied and the proof is complete. The optimal solution to the exit problem in this case divides the state space into two regions as shown in Figure 2. □

We next provide a sufficient condition in terms of the entries of A for the standard case to occur when $\hat{x}^i \in (\overline{\mathscr{B}}^i_{jk})^\circ$.

Lemma 6.3. *Let A be a simple three-strategy coordination game. Suppose $\hat{x}^i \in (\overline{\mathscr{B}}^i_{jk})^\circ$ i.e., $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$. A sufficient condition for the standard case i.e., for $DW^k(\hat{x}^i)(e_i - e_j) \ge 0$ to hold is*

$$3A_{i-j}^{i-k}A_{i-k}^{i-k} \geq A_{i-j}^{i+k-2j}A_{i-k}^{2j-i-k} \tag{6.7}$$

Proof. From (6.5), we have

$$\begin{aligned}
c(\tilde{x}_i^{ik}) &= DW^k(\tilde{x}_i^{ik})(e_i - e_j) \\
&= \frac{1}{4}\frac{(A^{i-k}\tilde{x}^{ik})^2}{A_{i-k}^{i-k}}A_{i-j}^{i-k} + \frac{1}{12}\frac{(A^{2j-i-k}\tilde{x}^{ik})^2}{A_{i-k}^{2j-i-k}}A_{i-j}^{2j-i-k} \\
&= \frac{1}{4}\frac{A_{i-j}^{i-k}}{A_{i-k}^{i-k}}(A^{i-k}\tilde{x}^{ik})^2 - \frac{1}{12}\frac{A_{i-j}^{i+k-2j}}{A_{i-k}^{2j-i-k}}(A^{j-k}\tilde{x}^{ik} - A^{i-j}\tilde{x}^{ik})^2 \\
&= \frac{1}{4}\frac{A_{i-j}^{i-k}}{A_{i-k}^{i-k}}(x_k^*A_{i-k}^{i-k})^2 - \frac{1}{12}\frac{A_{i-j}^{i+k-2j}}{A_{i-k}^{2j-i-k}}(x_k^*A_{i-k}^{j-k} - x_k^*A_{i-k}^{i-j})^2 \qquad \text{(since } \tilde{x}^{ik} = x^* + x_k^*(e_i - e_k)) \\
&= \frac{(x_k^*)^2}{12}\left(3\frac{A_{i-j}^{i-k}}{A_{i-k}^{i-k}}(A_{i-k}^{i-k})^2 - \frac{A_{i-j}^{i+k-2j}}{A_{i-k}^{2j-i-k}}(A_{i-k}^{j-k} - A_{i-k}^{i-j})^2\right) \\
&= \frac{(x_k^*)^2}{12}\left(3A_{i-j}^{i-k}A_{i-k}^{i-k} - A_{i-j}^{i+k-2j}A_{i-k}^{2j-i-k}\right) \qquad (6.8)
\end{aligned}$$

In Theorem 6.2, we have shown that the function $c(s) = DW^k(x(s))(e_i - e_j)$ for $x(s) = se_i + (1-s)e_j$, $\tilde{x}_i^{ik} \leq s \leq \underline{x}_i^{ij}$ is non-decreasing in s. Since $\hat{x}_i^i > \tilde{x}_i^{ik}$, a sufficient condition for $DW^k(\hat{x}^i)(e_i - e_j) \geq 0$ i.e., $c(\hat{x}_i^i) \geq 0$ to hold is $c(\tilde{x}_i^{ik}) \geq 0$ and therefore (6.7) follows from (6.8). ◻

6.2 Retreating Case

In what follows, we consider simple three strategy coordination games with $z^{ij} \in (\overline{\mathcal{B}}_{jk}^i)^\circ$ and $DW^k(z^{ij})(e_i - e_j) < 0$. We firstly note from Lemma 6.3, that a necessary condition for the retreating case is given by the following:

Corollary 6.4. *Let A be a simple three-strategy coordination game. Suppose $\hat{x}^i$ is in $(\overline{\mathcal{B}}_{jk}^i)^\circ$ i.e., $A_{i-k}^{i-k} \geq 8A_{i-j}^{i-j}$. A necessary condition for the retreating case i.e., for $DW^k(\hat{x}^i)(e_i - e_j) < 0$ to hold is*

$$3A_{i-j}^{i-k}A_{i-k}^{i-k} < A_{i-j}^{i+k-2j}A_{i-k}^{2j-i-k} \tag{6.9}$$

In the retreating case, we show that there exists a unique state $\overline{x}^i$ on $l^{ij} \cap \overline{\mathcal{B}}_{jk}^i$ with $\overline{x}_i^i > z_i^{ij}$ such that $DW^k(\overline{x}^i)(e_i - e_j) = 0$. We also show that there exists a unique state $\underline{x}^i$ on $l^{ij} \cap \overline{\mathcal{B}}_{jk}^i$ with $\underline{x}_i^i < z_i^{ij}$ such that $W^j(\underline{x}^i) = W^k(\overline{x}^i)$. We prove the existence and uniqueness of these states in Lemma 6.5. These two states play an important role in the retreating case in describing the optimal solution. In Theorem 6.7, we show that when $\overline{x}^i_i \le 1$, the optimal exit path will not have binding state constraints. In Theorem 6.8, we show that when $\overline{x}^i_i > 1$, the optimal exit path has binding state constraints.

Lemma 6.5. *If $z^{ij} \in \overline{\mathcal{B}}^i_{jk}$ and $DW^k(z^{ij})(e_i - e_j) < 0$, then there is a unique state $\overline{x}^i$ on $l^{ij} \cap \overline{\mathcal{B}}^i_{jk}$ with $z^{ij}_i < \overline{x}^i_i \le \underline{x}^{ij}_i$ such that $DW^k(\overline{x}^i)(e_i - e_j) = 0$ and a unique state $\underline{x}^i$ on l^{ij} with $\tilde{x}^{ik}_i < \underline{x}^i_i < z^{ij}_i$ such that $W^j(\underline{x}^i) = W^k(\overline{x}^i)$.*

Proof. See Appendix G. □

We need some notation before stating our main results for the retreating case. Let

$$w^i_1 = x^* \times (\underline{x}^i - x^*)$$
$$w^i_2 = x^* \times (\overline{x}^i - x^*)$$

Under the assumptions in Lemma 6.5, a homogeneity argument leads to the following:

Corollary 6.6. *For any state $x \in \mathcal{B}^i$ with $(w^i_1)'x < 0$ and $(w^i_2)'x > 0$, there exists a unique state $\overline{x}$ in $\overline{\mathcal{B}}^i_{jk}$ with $\overline{x}_i > x_i$ such that the line joining x and $\overline{x}$ is parallel to $e_i - e_j$ and $DW^k(\overline{x})(e_i - e_j) = 0$. There exists a unique state $\underline{x}$ in $\overline{\mathcal{B}}^i_{jk}$ with $\underline{x}_i < x_i$ such that the line joining x and $\underline{x}$ is parallel to $e_i - e_j$ and $W^j(\underline{x}) = W^k(\overline{x})$.*

Table 3: Important States in Retreating Case

Point	Description
$\overline{x}^i$	unique state on $l^{ij} \cap \overline{\mathcal{B}}^i_{jk}$ such that $DW^k(\overline{x}^i)(e_i - e_j) = 0$
$\underline{x}^i$	unique state on $l^{ij} \cap \overline{\mathcal{B}}^i_{jk}$ such that $W^j(\underline{x}^i) = W^k(\overline{x}^i)$

6.2.1 Retreating without binding state constraints

Theorem 6.7. *Let A be a simple three-strategy coordination game such that $\hat{x}^i \in (\overline{\mathcal{B}}^i_{jk})^\circ$ i.e., $A^{i-k}_{i-k} > 8A^{i-j}_{i-j}$. Suppose that $DW^k(\hat{x}^i)(e_i - e_j) < 0$. Further suppose that $\overline{x}^i_i \le 1$. In this case, the value function $W^* : \mathcal{B}^i \to \mathbb{R}_+$ for the exit cost problem with target set $\mathcal{B}^j \cup \mathcal{B}^k$ is given by the continuous function*

$$
(6.10) \qquad W^*(x) = \begin{cases} = W^k(x) & \text{if } (w_2^i)'x \le 0 \\ = W^k(\overline{x}) & \text{if } (w_1^i)'x < 0 \text{ and } (w_2^i)'x > 0, \\ = W^j(x) & \text{if } (w_1^i)'x \ge 0. \end{cases}
$$

The optimal feedback controls in this case are (see Figure 6)

$$
(6.11) \qquad \nu^*(x) = \begin{cases} = e_k - e_i & \text{if } (w_2^i)'x \le 0 \\ = e_i - e_j & \text{if } (w_1^i)'x < 0 \text{ and } (w_2^i)'x > 0, \\ \in \{e_i - e_j, e_j - e_i\} & \text{if } (w_1^i)'x = 0, \\ = e_j - e_i & \text{if } (w_1^i)'x > 0. \end{cases}
$$

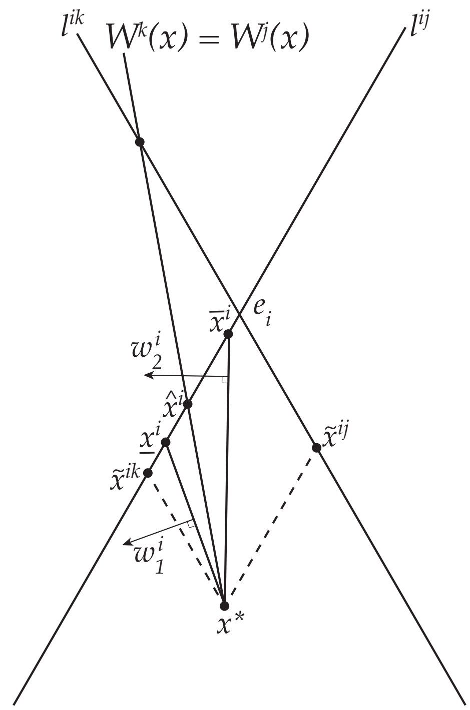


Figure 6: Retreating with non-binding state constraints

Proof. From Lemma 6.5, $z_i^{ij} < \overline{x}_i^i$. Since $\overline{x}_i^i \le 1$ by assumption, we have $z_i^{ij} < 1$. This implies that $\hat{x}^i$ is on the line l^{ij}. Following the arguments in the proofs of Theorems 6.1 and 6.2, it is sufficient to show that

(a) For x on l^{ij} with $\underline{x}_i^i < x_i < \overline{x}_i^i$ moving in the direction $e_i - e_j$ until reaching $\overline{x}^i$ and then moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ satisfies the HJB equation.

(b) For x on l^{ij} with $\overline{x}_i^i < x_i < x_i^{ijk}$ moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ satisfies the HJB equation.

Consider the states x in (b) with $\overline{x}_i^i < x_i \le \underline{x}_i^{ij}$. We have shown in the proof of Theorem 6.2 that $DW^k(x)(e_i - e_j)$ is non-decreasing in x_i for states x with $\tilde{x}_i^{ik} \le x_i \le \underline{x}_i^{ij}$. For the states under consideration, the function $DW^k(x)(e_i - e_j)$ therefore attains its minimum at $\overline{x}^i$. By definition, we have $DW^k(\overline{x}^i)(e_i - e_j) = 0$. This implies that $DW^k(x)(e_i - e_j) \ge 0$ for $\overline{x}_i^i \le x_i \le \underline{x}_i^{ij}$. From the arguments in the proof of Theorem 6.2, it follows that the HJB equation is satisfied for such states. Now consider the states x in (b) with $\underline{x}_i^{ij} \le x_i < x_i^{ijk}$. For such x, by definition, we have $x \in \underline{\mathcal{B}}_{jk}^i$. From computations in Appendix E, it follows that the HJB equation is satisfied for such states. Therefore, the HJB equation is satisfied for states x in (b). We verify (a) in Appendix H.

We now apply the verification theorem. The value function W^* in (6.10) is constructed from feedback controls (6.11) that generate feasible solutions to the exit problem, as required by condition (i) of Theorem 5.1. The continuity of W^* follows from the construction of the points $\overline{x}^i$ and $\underline{x}^i$. W^* is clearly C^1 off the sets $\{x \in \mathrm{aff}(X) : (w_1^i)'x = 0\}$ and $\{x \in \mathrm{aff}(X) : (w_2^i)'x = 0\}$. The arguments above imply that the HJB equation holds *a.e.* away from these sets. Thus condition (ii) of Theorem 5.1 is satisfied and the proof for the *retreating case* without binding state constraints is complete. The optimal solution to the exit problem in this case divides the state space into three regions as shown in Figure 3. □

Theorem 6.7 says that in the retreating case, a sufficient condition for the optimal exit path to have non-binding state constraints is $\overline{x}_i^i \le 1$. In Section 6.2.2, we show that $\overline{x}_i^i \le 1$ is also a necessary condition to have non-binding state constraints i.e., when $\overline{x}_i^i > 1$, the optimal exit path will have binding state constraints.

6.2.2 Retreating with binding state constraints

We need some more notation before presenting the main result for the retreating case with binding state constraints. Suppose $\overline{x}_i^i > 1$. Let the line joining x^* and $\overline{x}^i$ intersect l^{ik} at the point $\overline{z}^i$ (see Figure 7). From Lemma 6.5 and a homogeneity argument it follows that there is a unique state $\underline{z}^i$ on the line parallel to $e_j - e_i$ through $\overline{z}^i$ such that $\underline{z}_i^i < \overline{z}_i^i$ and $W^j(\underline{z}^i) = W^k(\overline{z}^i)$.

If $z_i^{ij} \le 1$ i.e., $\hat{x}^i \in l^{ij} \cap X$, then there exists a unique state $\underline{y}^i$ on the line l^{ij} with $\tilde{x}_i^{ik} < \underline{y}_i^i < \hat{x}_i^i$ such that $W^j(\underline{y}^i) = W^k(e_i)$. Existence and uniqueness of $\underline{y}^i$ follows from Lemma 6.5.[4] Let z be any state on l^{ik} between the states e_i and $\overline{z}^i$ i.e., $\overline{z}_i^i \le z_i \le e_i$. Lemma 6.5 and a homogeneity argument implies that there exists a unique state $y \in \overline{\mathcal{B}}_{jk}^i$ on the line parallel to $e_j - e_i$ passing

[4] A slight modification is needed in defining the function h in the proof of Lemma 6.5. Defining the function as $h(s) = W^j(x(s)) - W^k(e_i)$ and then following the same steps as in the proof of Lemma 6.5, proves the existence of $\underline{y}^i$.

through z with $y_i < z_i$ such that $W^j(y) = W^k(z)$. The collection of these points y define a continuous curve (denote this by γ_1) in $\mathcal{B}^i$ as the functions $W^j(x)$ and $W^k(x)$ are continuous. By construction this continuous curve has $\underline{y}^i$ and $\underline{z}^i$ as the endpoints. Let R_1 be the region enclosed by the lines $\underline{y}^i x^{ij}$, $x^{ij}x^*$, $x^*\underline{z}^i$ and the continuous curve γ_1. Let R_2 denote the region enclosed by the continuous curve γ_1 and the lines $\underline{z}^i x^*$, $x^*\overline{z}^i$, $\overline{z}^i e_i$, $e_i\underline{y}^i$. We further subdivide R_2 into R_2^1 and R_2^2 as follows : R_2^1 is the region enclosed by the continuous curve γ_1 and the lines $\underline{z}^i\overline{z}^i$, $\overline{z}^i e_i$, $e_i\underline{y}_i$. R_2^2 is the convex hull of the states $\underline{z}^i$, x^* and $\overline{z}^i$. Let R_3 be the convex hull of the points $\overline{z}^i$, x^* and x^{ik} (see Figure 7(i)).

If $z_i^{ij} > 1$, i.e., $\hat{x}^i \in l^{ik} \cap X$, then Lemma 6.5 and a homogeneity argument again imply that there exists a continuous curve (denote this by γ_2) joining $\underline{z}^i$ and $\hat{x}_i^i$ such that for any point y on this curve, we have $W^j(y) = W^k(z)$ where z is on l^{ik} with $\overline{z}_i^i \le z_i \le \hat{x}_i^i$ and yz is parallel to e_ie_j. Let R_1 be the region enclosed by the lines e_ix^{ij}, $x^{ij}x^*$, $x^*\underline{z}^i$ and the continuous curve γ_2. Let R_2 denote the region enclosed by the continuous curve γ_2 and the lines $\underline{z}^i x^*$, $x^*\overline{z}^i$, $\overline{z}^i\hat{x}^i$. We further subdivide R_2 into R_2^1 and R_2^2 as follows : R_2^1 is the region enclosed by the continuous curve γ_2 and the lines $\underline{z}^i\overline{z}^i$, $\overline{z}^i\hat{x}^i$. R_2^2 is the convex hull of the states $\underline{z}^i$, x^* and $\overline{z}^i$. Let R_3 be the convex hull of the states $\overline{z}^i$, x^* and x^{ik} (see Figure 7(ii)).

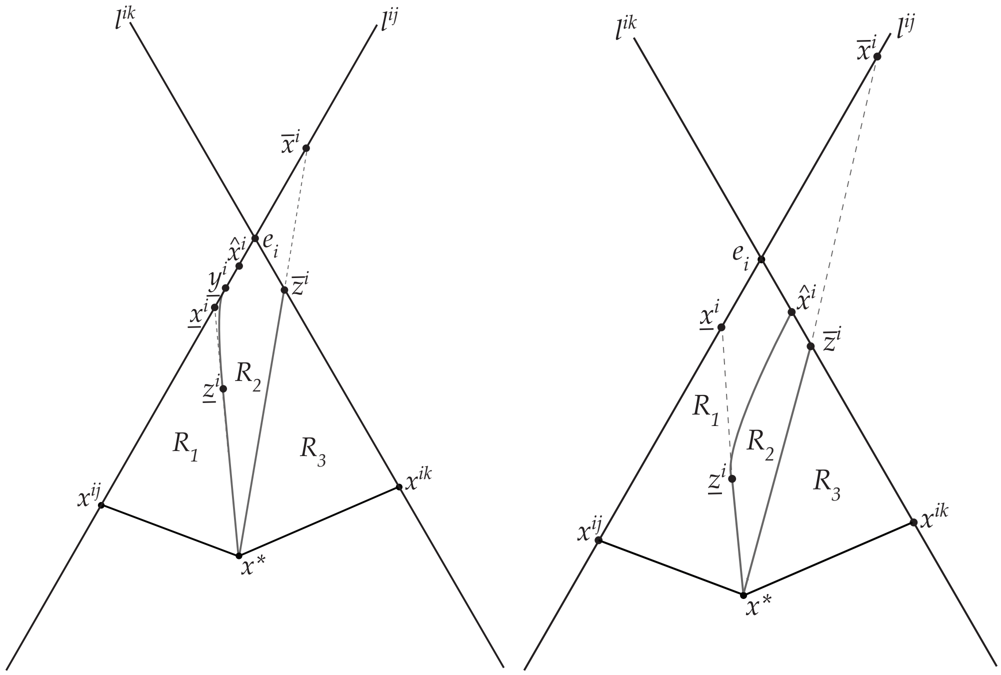
For $x \in R_2^1$, we denote by $z(x)$ the unique state on l^{ik} such that the line joining $z(x)$ and x is parallel to $e_i - e_j$. For $x \in R_2^2$, we denote by $z(x)$ the unique state on the line $x^*\overline{z}^i$ such that the line joining $z(x)$ and x is parallel to $e_i - e_j$.

Theorem 6.8. *Let A be a simple three-strategy coordination game such that $\hat{x}^i \in (\overline{\mathcal{B}}_{jk}^i)^\circ$ i.e., $A_{i-k}^{i-k} > 8A_{i-j}^{i-j}$. Suppose that $DW^k(\hat{x}^i)(e_i - e_j) < 0$. Further suppose that $\overline{x}_i^i > 1$. In this case, the value function $W^* : \mathcal{B}^i \to \mathbb{R}_+$ for the exit cost problem with target set $\mathcal{B}^j \cup \mathcal{B}^k$ is given by the continuous function*

$$W^*(x) = \begin{cases} = W^k(x) & \text{if } x \in R_3 \\ = W^k(z(x)) & \text{if } x \in R_2, \\ = W^j(x) & \text{if } x \in R_1. \end{cases} \tag{6.12}$$

The optimal feedback controls in this case are (see Figure 7)

$$\nu^*(x) = \begin{cases} = e_k - e_i & \text{if } x \in R_3 \\ = e_i - e_j & \text{if } x \in R_2^\circ, \\ \in \{e_i - e_j, e_j - e_i\} & \text{if } x \in bd(R_1 \cap R_2), \\ = e_j - e_i & \text{if } x \in R_1. \end{cases} \tag{6.13}$$



(i) $\hat{x}^i$ is in $\overline{\mathcal{B}}^i_{jk}$ and on the face e_ie_j (ii) $\hat{x}^i$ is in $\overline{\mathcal{B}}^i_{jk}$ and on the face e_ie_k

Figure 7: Retreating with binding state constraints

Proof. We need to show that

(a) For states x in $(R_2^1)^\circ$ moving in the direction $e_i - e_j$ until reaching the boundary of the simplex l^{ik} and then moving along the boundary in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ satisfies the HJB equation.

(b) For states x in $(R_2^2)^\circ$ moving in the direction $e_i - e_j$ until reaching the line joining the points $\overline{z}^i$ and x^* and then moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ satisfies the HJB equation.

(c) For states x on the line l^{ik} with $x_i^{ikj} < x_i < \overline{z}_i^i$ moving in the direction $e_k - e_i$ until reaching $\mathcal{B}^{ik}$ satisfies the HJB equation.

Homogeneity argument (6.2) implies (c) is equivalent to proving that for the states x on the line l^{ij} with $\overline{x}_i^i < x_i < x_i^{ijk}$ moving in the direction $e_k - e_i$ is consistent with the HJB equation. But this follows directly from (b) in the proof of Theorem 6.7. Another application of homogeneity argument (6.2) shows that (b) is equivalent to proving that the HJB equation is satisfied for states x on l^{ij} with $\underline{x}_i^i < x_i < \overline{x}_i^i$. But this follows directly from (a) in the proof of Theorem 6.7. We prove (a) in Appendix I.

We now apply the verification theorem. The value function W^* in (6.12) is constructed from feedback controls (6.13) that generate feasible solutions to the exit problem, as required by condition (i) of Theorem 5.1. The continuity of W^* follows from construction.

The arguments above imply that the HJB equation holds *a.e.* in the interior of X. Thus condition (ii) of Theorem 5.1 is satisfied and the proof is complete. The optimal solution to the exit problem in this case divides the state space into three regions as shown in Figure 4. $\square$

7. Conclusion

In this paper we considered the exit problem, a control problem associated with large deviations properties, which is used to assess the expected time until the evolutionary process leaves the basin of attraction of a stable equilibrium and to determine the likely exit path. Solving this problem for simple three strategy coordination games under probit choice rule, we show that the likely exit path can be as shown in Figures 2 to 4 depending on the strength of the coordination incentives between the equilibrium strategy and the alternative strategies.

In order to evaluate stationary distribution asymptotics and stochastic stability, one must consider the transition problem, a control problem associated with large deviations properties, which is used to assess the probable time until a transition between a given pair of stable equilibria and to determine the most likely path that this transition will follow. The exit problem involves finding the least cost paths from the initial basin of attraction to all the states in the state space, whereas the transition problem involves finding the least cost paths from all the states in the state space to the initial basin of attraction. Since the form of the unlikelihood function changes as we traverse the state space, computing all the feasible path costs and then finding the optimal path in the transition problem is more involved than finding the optimal path in the exit problem. However, our analysis to solve the exit problem provides a good starting point to split the transition problem into several cases based on the strength of coordination incentives and then solve each of this case separately. We leave this for future research.

Appendix

A. Unlikelihood function for three-strategy probit choice model in the best response region of strategy *i*

For any set $K \subseteq S$ of cardinality n^K, let

$$\overline{\pi}^K = \frac{1}{n^K}\sum_{k\in K}\pi_k$$

denote the average payoff of the actions in K. For $l \in \{1, 2, \cdots, n\}$, Dokumacı and Sandholm (2011) show that the unlikelihood function Υ is given by

$$\Upsilon_l(\pi) = \sum_{a=1}^{n}\frac{(z_a^*)^2}{2}, \text{ where}$$

$$z_b^* = \begin{cases}\overline{\pi}^{J\cup\{l\}} - \pi_b & if\ b \in J \cup \{l\}\\ 0 & otherwise,\end{cases}$$

with the set $J \subset S - \{l\}$ being uniquely determined by the requirement that

$$j \in J\ if\ and\ only\ if\ \pi_j > \overline{\pi}^{J\cup\{l\}}$$

Consider a strategy set $S = \{i, j, k\}$. We use the above result to derive the unlikelihood function Υ in the best response region of strategy i. Let $\pi = (\pi_i, \pi_j, \pi_k)$. We first compute Υ_i. Since $\pi_i \geq \max\{\pi_j, \pi_k\}$, we have $J = \emptyset$. Therefore, we have $z_l^* = 0$ for $l = 1, 2, 3$. Hence, $\Upsilon_i(\pi) = 0$.

Next, we compute Υ_j. We consider the following exhaustive cases:

Case 1: $\pi_k \leq \frac{\pi_i+\pi_j}{2}$

In this case $J = \{i\}$. Therefore, we have $z_i^* = \frac{1}{2}(\pi_j - \pi_i)$, $z_j^* = \frac{1}{2}(\pi_i - \pi_j)$ and $z_k^* = 0$. Hence,

$$\Upsilon_j(\pi) = \frac{1}{4}(\pi_i - \pi_j)^2$$

Case 2: $\pi_k \geq \frac{\pi_i+\pi_j}{2}$

In this case $J = \{i, k\}$. Therefore, we have $z_i^* = \frac{1}{3}(\pi_k + \pi_j - 2\pi_i)$, $z_j^* = \frac{1}{3}(\pi_i + \pi_k - 2\pi_j)$ and $z_k^* = \frac{1}{3}(\pi_i + \pi_j - 2\pi_k)$. A simple computation gives us

$$\Upsilon_j(\pi) = \frac{1}{6}\Big((\pi_j - \pi_i)^2 + (\pi_k - \pi_j)^2 + (\pi_i - \pi_k)^2\Big)$$

By interchanging j and k in the above computations we get $\Upsilon_k(\pi)$.

B. Proof of Lemma 5.3

We first show that $A_{i-k}^{2j-i-k} > 0$. Let $x \in \overline{\mathcal{B}}_{jk}^i$, with $x_j < x_j^*$. We can assume without loss of generality that this point x is on the line l^{ij} i.e., $x(s) = se_i + (1-s)e_j$ for some $s \in \mathbb{R}$. Define a function $z(s)$ as follows

$$z(s) = A^{j-k}x(s) - A^{i-j}x(s)$$

A simple computation gives us

$$\begin{aligned} z'(s) &= A_{i-j}^{j-k} - A_{i-j}^{i-j} \\ &= -(A_{i-j}^{k-j} + A_{i-j}^{i-j}) \\ &< 0 && \text{(by CG (4.3) and MBP (4.4))} \end{aligned}$$

Since $\tilde{x}^{ik} = x^* + x_k^*(e_i - e_k)$, we have $\tilde{x}_i^{ik} = x_i^* + x_k^*$ and $\tilde{x}_j^{ik} = x_j^*$.

$$\begin{aligned} z(\tilde{x}_i^{ik}) &= A^{j-k}\tilde{x}^{ik} - A^{i-j}\tilde{x}^{ik} \\ &= x_k^*(A_{i-k}^{j-k} - A_{i-k}^{i-j}) \\ z(x_i) &= A^{j-k}x(s) - A^{i-j}x(s) \\ &\geq 0 && \text{(since } x(s) \in \overline{\mathcal{B}}_{jk}^{i}) \end{aligned}$$

Under our hypothesis $x_j < x_j^*$ i.e., $1 - x_j > 1 - x_j^*$. This implies that $x_i > \tilde{x}_i^{ik}$. As $z'(s) < 0$, we have $z(\tilde{x}_i^{ik}) > z(x_i)$. From the above expressions, we therefore have $A_{i-k}^{j-k} - A_{i-k}^{i-j} > 0$ i.e., $A_{i-k}^{2j-i-k} > 0$.

We now compute the path costs. For $x \in \overline{\mathcal{B}}_{jk}^{i}$, there is a $t_0 \in (0,1)$ such that the path ϕ_t enters the region $\underline{\mathcal{B}}_{jk}^{i}$ for the first time (see Figure 8) i.e., $\phi_t \in \overline{\mathcal{B}}_{jk}^{i}$ for $t < t_0$, $\phi_t \in \underline{\mathcal{B}}_{jk}^{i}$ for $t > t_0$ and at $t = t_0$ it is on the boundary of $\overline{\mathcal{B}}_{jk}^{i}$ and $\underline{\mathcal{B}}_{jk}^{i}$. From Lemma 2.3, we have $e_k'\Upsilon(F(\phi_t)) = \frac{1}{6}\left[(F_i(\phi_t) - F_j(\phi_t))^2 + (F_j(\phi_t) - F_k(\phi_t))^2 + (F_i(\phi_t) - F_k(\phi_t))^2\right]$, when $\phi_t \in \overline{\mathcal{B}}_{jk}^{i}$. For $\phi_t \in \underline{\mathcal{B}}_{jk}^{i}$, $e_k'\Upsilon(F(\phi_t)) = \frac{1}{4}(F_i(\phi_t) - F_k(\phi_t))^2$.

We now compute as follows [5]

$$\begin{aligned} \gamma(x,y) &= \int_0^1 [\dot{\phi}_t]_+'\Upsilon(F(\phi_t))dt \\ &= \int_0^1 de_k'\Upsilon(F(\phi_t))dt \\ &= \int_0^{t_0} de_k'\Upsilon(F(\phi_t))dt + \int_{t_0}^1 de_k'\Upsilon(F(\phi_t))dt \\ &= \int_0^{t_0} \frac{d}{6}\left[(\pi_i - \pi_j)^2 + (\pi_j - \pi_k)^2 + (\pi_i - \pi_k)^2\right]dt + \int_{t_0}^1 \frac{d}{4}(\pi_i - \pi_k)^2 dt \\ &= \int_0^{t_0} \frac{d}{6}\left[(\pi_i - \pi_j)^2 + (\pi_j - \pi_k)^2 + (\pi_i - \pi_k)^2\right]dt - \int_0^{t_0} \frac{d}{4}(\pi_i - \pi_k)^2 dt \end{aligned}$$

[5] Recall $\pi_l = F_l(\phi_t)$, for $l = \{i, j, k\}$

$$+\int_0^{t_0}\frac{d}{4}(\pi_i-\pi_k)^2dt+\int_{t_0}^1\frac{d}{4}(\pi_i-\pi_k)^2dt$$

$$=\int_0^{t_0}\frac{d}{6}\left[(\pi_i-\pi_j)^2+(\pi_j-\pi_k)^2-\frac{1}{2}(\pi_i-\pi_k)^2\right]dt+\int_0^1\frac{d}{4}(\pi_i-\pi_k)^2dt$$

$$=c(x)+\frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} \tag{B.1}$$

where $c(x)=\int_0^{t_0}\frac{d}{6}\left[(\pi_i-\pi_j)^2+(\pi_j-\pi_k)^2-\frac{1}{2}(\pi_i-\pi_k)^2\right]dt$.

$\int_0^1\frac{d}{4}(\pi_i-\pi_k)^2dt=\frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}}$ follows from computations in Lemma 5.2 by replacing the index j with k.

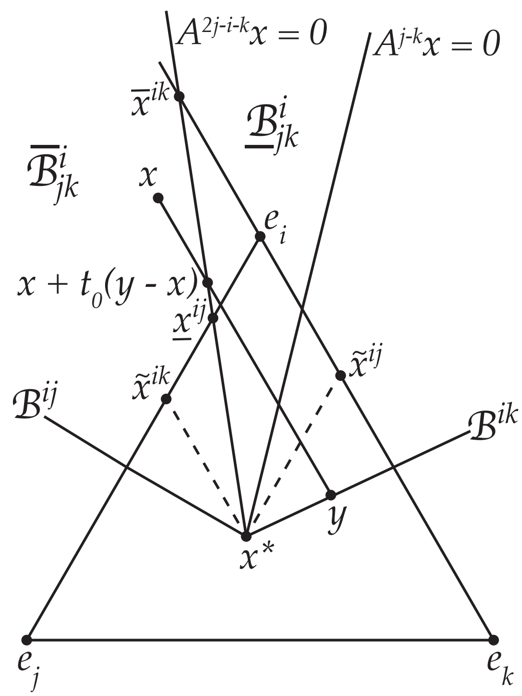


Figure 8: Correction Factor

We next compute the value of t_0. By definition, $\phi(t_0)$ is on the line $\pi_k+\pi_i=2\pi_j$ i.e., $(1-t_0)x+t_0y$ is on the line $\pi_i-\pi_j=\pi_j-\pi_k$. Thus t_0 solves the equation $(1-t_0)A^{i-j}x+t_0A^{i-j}y=(1-t_0)A^{j-k}x+t_0A^{j-k}y$. Using the fact that $y=x+d(e_k-e_i)$, and simplifying the above equation, we get

$$t_0=\frac{A^{j-k}x-A^{i-j}x}{d(A^{j-k}_{i-k}-A^{i-j}_{i-k})}=\frac{A^{2j-i-k}x}{dA^{2j-i-k}_{i-k}} \tag{B.2}$$

By definition, we have

$$(\pi_j-\pi_k)-(\pi_i-\pi_j)=(1-t)A^{j-k}x+tA^{j-k}y-(1-t)A^{i-j}x-tA^{i-j}y$$

$$\begin{aligned}
&= (A^{j-k}x - A^{i-j}x) + t(A^{j-k}y - A^{j-k}x) - t(A^{i-j}y - A^{i-j}x) \\
&= (A^{j-k}x - A^{i-j}x) - tdA^{j-k}_{i-k} + tdA^{i-j}_{i-k} \\
&= d(A^{j-k}_{i-k} - A^{i-j}_{i-k})\left(\frac{(A^{j-k}x - A^{i-j}x)}{d(A^{j-k}_{i-k} - A^{i-j}_{i-k})} - t\right) \qquad \text{(B.3)} \\
&= dA^{2j-i-k}_{i-k}(t_0 - t) \qquad \text{(from (B.2))}
\end{aligned}$$

We know for any $a, b, c \in \mathbb{R}$,

$$(a-b)^2 + (b-c)^2 - \frac{1}{2}(c-a)^2 = \frac{1}{2}((b-c)-(a-b))^2 \tag{B.4}$$

We now compute the correction term $c(x)$ from (B.1) as follows

$$\begin{aligned}
c(x) &= \int_0^{t_0} \frac{d}{6}\left[(\pi_i - \pi_j)^2 + (\pi_j - \pi_k)^2 - \frac{1}{2}(\pi_i - \pi_k)^2\right] dt \\
&= \frac{d}{12}\int_0^{t_0} ((\pi_j - \pi_k) - (\pi_i - \pi_j))^2 dt \qquad \text{(from (B.4))} \\
&= \frac{d}{12}\int_0^{t_0} (dA^{2j-i-k}_{i-k}(t_0 - t))^2 dt \qquad \text{(from (B.3))} \\
&= \frac{d^3(A^{2j-i-k}_{i-k})^2}{12}\int_0^{t_0} (t_0 - t)^2 dt \\
&= \frac{d^3(A^{2j-i-k}_{i-k})^2}{12}\frac{t_0^3}{3} \\
&= \frac{(A^{2j-i-k}_{i-k})^2}{36}(t_0 d)^3 \\
&= \frac{(A^{2j-i-k}_{i-k})^2}{36}\left(\frac{A^{2j-i-k}x}{A^{2j-i-k}_{i-k}}\right)^3 \qquad \text{(from (B.2))} \\
&= \frac{1}{36}\frac{(A^{2j-i-k}x)^3}{A^{2j-i-k}_{i-k}}
\end{aligned}$$

C. Proof of Lemma 5.5

We first verify the HJB equation (5.2) for $x \in (\overline{\mathcal{B}}^i_{jk})^\circ$.

$$\begin{aligned}
H(x, e_j - e_i) &= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}}A^{i-j}_{j-i} \\
&= 0
\end{aligned}$$

$$
\begin{aligned}
H(x, e_i - e_j) &= 0 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}} A^{i-j}_{i-j} \\
&\geq 0 \\
H(x, e_i - e_k) &= 0 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}} A^{i-j}_{i-k} \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))} \\
H(x, e_k - e_j) &= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{k-i}x)^2\right) + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}} A^{i-j}_{k-j} \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))} \\
H(x, e_j - e_k) &= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}} A^{i-j}_{j-k} \\
&= \frac{1}{4}(A^{i-j}x)^2 \left(\frac{A^{i-j}_{i-k}}{A^{i-j}_{i-j}}\right) \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))}
\end{aligned}
$$

For $x \in \overline{\mathcal{B}}^i_{jk}$, by definition $A^k x + A^i x \leq 2A^j x$ i.e., $A^{i-j}x < A^{j-k}x$. We now compute as follows

$$
\begin{aligned}
H(x, e_k - e_i) &= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2\right) + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}} A^{i-j}_{k-i} \\
&= \frac{1}{6}\left((A^{j-k}x)^2 + (A^{i-k}x)^2\right) + \frac{1}{4}\left(1 - \frac{A^{i-j}_{i-k}}{A^{i-j}_{i-j}}\right)(A^{i-j}x)^2 - \frac{1}{12}(A^{i-j}x)^2 \\
&= \frac{1}{6}\left((A^{j-k}x)^2 + (A^{i-k}x)^2\right) + \frac{1}{4}\left(\frac{A^{i-j}_{k-j}}{A^{i-j}_{i-j}}\right)(A^{i-j}x)^2 - \frac{1}{12}(A^{i-j}x)^2 \\
&\geq \frac{1}{6}\left((A^{j-k}x)^2 + (A^{i-k}x)^2\right) - \frac{1}{12}(A^{i-j}x)^2 && \text{(by CG (4.3) and MBP (4.4))} \\
&\geq \frac{1}{12}\left((A^{j-k}x)^2 - (A^{i-j}x)^2\right) + \frac{1}{6}(A^{i-k}x)^2 \\
&\geq 0 && \text{(since } A^{j-k}x \geq A^{i-j}x\text{)}
\end{aligned}
$$

We next verify the HJB equation (5.2) for $x \in (\underline{\mathcal{B}}^i_{jk})^\circ$. It is easily verified that the computations in the directions of $e_j - e_i$, $e_i - e_j$, $e_i - e_k$ and $e_j - e_k$ are exactly the same as $x \in (\overline{\mathcal{B}}^i_{jk})^\circ$. For the remaining two directions $e_k - e_j$ and $e_k - e_i$:

$$
H(x, e_k - e_j) = \frac{1}{4}(A^{i-k}x)^2 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}} A^{i-j}_{k-j}
$$

$$\geq 0 \qquad \text{(by CG (4.3) and MBP (4.4))}$$

We have $A^j x \geq A^k x$. This implies that $A^{i-k}x \geq A^{i-j}x$ and so we have,

$$\begin{aligned}
H(x, e_k - e_i) &= \frac{1}{4}(A^{i-k}x)^2 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}}A^{i-j}_{k-i} \\
&\geq \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-j}x)^2}{A^{i-j}_{i-j}}A^{i-j}_{k-i} \\
&= \frac{1}{4}(A^{i-j}x)^2\left(\frac{A^{i-j}_{k-j}}{A^{i-j}_{i-j}}\right) \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))}
\end{aligned}$$

D. Proof of Theorem 5.7

For $x(s) = se_i + (1-s)e_j,\ s \in \mathbb{R}$, let

(D.1) $\qquad g(s) = W^k(x(s)) - W^j(x(s))$ for $s \geq \tilde{x}_i^{ik}$

It suffices to show that the equation $g(s) = 0$ has a unique real root for $s > \tilde{x}_i^{ik}$. Let

(D.2) $$f(s) = \frac{1}{12}\left(\frac{(A^{i-k}x(s))^3}{A^{i-k}_{i-k}} - \frac{(A^{i-j}x(s))^3}{A^{i-j}_{i-j}}\right)$$

$$a_1(s) = \begin{cases} \frac{1}{36}\frac{(A^{2j-i-k}x(s))^3}{A^{2j-i-k}_{i-k}} & \text{if } x(s) \in \overline{\mathcal{B}}^i_{jk} \text{ and } s \geq \tilde{x}_i^{ik} \\ 0 & \text{otherwise} \end{cases}$$

$$a_2(s) = \begin{cases} -\frac{1}{36}\frac{(A^{2k-i-j}x(s))^3}{A^{2k-i-j}_{i-j}} & \text{if } x(s) \in \overline{\mathcal{B}}^i_{kj} \text{ and } s \geq \tilde{x}_i^{ik} \\ 0 & \text{otherwise} \end{cases}$$

From (5.6), (5.7), (D.1), (D.2) and the above set of equations, it follows that

(D.3) $\qquad g(s) = f(s) + a_1(s) + a_2(s)$ for $s \geq \tilde{x}_i^{ik}$.

We prove the theorem by establishing three Lemmas. In Lemma D.1, we show that $f(s)$ is a cubic polynomial which has a unique real root for $s > \tilde{x}_i^{ik}$. In Lemma D.3, we prove that $a_1(s)$ is a non-increasing non-negative convex function for $s \geq \tilde{x}_i^{ik}$. In Lemma D.4, we prove that $a_2(s)$ is a non-increasing non-positive concave function for $s \geq \tilde{x}_i^{ik}$.

Lemma D.1. *There is a unique zero of the function $f(s)$ for $s > \tilde{x}_i^{ik}$.*

Proof of Lemma D.1. We first show that $f(\tilde{x}_i^{ik}) > 0$. Note that

$$
\begin{aligned}
A^{i-k}\tilde{x}^{ik} &= A^{i-k}(x^* + x_k^*(e_i - e_k)) \\
&= A^{i-k}x^* + x_k^* A^{i-k}(e_i - e_k) \\
&= x_k^* A_{i-k}^{i-k} && \text{(since } A^{i-k}x^* = 0 \text{ as } x^* \text{ is a mixed equilibrium)}
\end{aligned}
$$

Similarly, we have $A^{i-j}\tilde{x}^{ik} = x_k^* A_{i-k}^{i-j}$. From (D.2), we have

$$
\begin{aligned}
f(\tilde{x}_i^{ik}) &= \frac{1}{12}\left(\frac{(A^{i-k}\tilde{x}^{ik})^3}{A_{i-k}^{i-k}} - \frac{(A^{i-j}\tilde{x}^{ik})^3}{A_{i-j}^{i-j}}\right) \\
&= \frac{(x_k^*)^3}{12}\left(\frac{(A_{i-k}^{i-k})^3}{A_{i-k}^{i-k}} - \frac{(A_{i-k}^{i-j})^3}{A_{i-j}^{i-j}}\right) \\
&= \frac{(x_k^*)^3}{12}\left(\frac{(A_{i-k}^{i-k})^2 A_{i-j}^{i-j} - (A_{i-k}^{i-j})^3}{A_{i-j}^{i-j}}\right) \\
&> 0 \qquad \text{(D.4)}
\end{aligned}
$$

The last inequality holds as $A_{i-k}^{i-k} > A_{i-k}^{i-j}$ and $A_{i-j}^{i-j} > A_{i-k}^{i-j}$ follow from MBP (4.4).

Clearly from (D.2), $f(s)$ is a cubic polynomial with leading coefficient $\frac{1}{12}\left[\frac{(A_{i-j}^{i-k})^3}{A_{i-k}^{i-k}} - (A_{i-j}^{i-j})^2\right]$.

$A_{i-j}^{i-k} < A_{i-k}^{i-k}$ and $A_{i-j}^{i-k} < A_{i-j}^{i-j}$ follow from MBP (4.4). From this it follows that the leading coefficient of $f(s)$ is negative.

Differentiating (D.2), gives us

$$
f'(s) = \frac{1}{4}\left(\frac{A_{i-j}^{i-k}(A^{i-k}x(s))^2}{A_{i-k}^{i-k}} - (A^{i-j}x(s))^2\right) \qquad \text{(D.5)}
$$

Differentiating (D.5), gives us

$$
f''(s) = \frac{1}{2}\left(\frac{(A_{i-j}^{i-k})^2 A^{i-k}x(s)}{A_{i-k}^{i-k}} - A_{i-j}^{i-j}A^{i-j}x(s)\right) \qquad \text{(D.6)}
$$

If $A_{i-j}^{i-k} \geq A_{i-k}^{i-j}$, then from (D.5)

$$
\begin{aligned}
f'(\tilde{x}_i^{ik}) &= \frac{1}{4}\left(\frac{A_{i-j}^{i-k}(A^{i-k}\tilde{x}^{ik})^2}{A_{i-k}^{i-k}} - (A^{i-j}\tilde{x}^{ik})^2\right) \\
&= \frac{x_k^{*2}}{4}\left(A_{i-j}^{i-k}A_{i-k}^{i-k} - (A_{i-k}^{i-j})^2\right) && \text{(since } \tilde{x}^{ik} = x^* + x_k^*(e_i - e_k)\text{)} \\
&\geq \frac{x_k^{*2}}{4}\left(A_{i-k}^{i-j}A_{i-k}^{i-k} - (A_{i-k}^{i-j})^2\right)
\end{aligned}
$$

$$= \frac{x_k^{*2}}{4} A_{i-k}^{i-j} \left(A_{i-k}^{i-k} - A_{i-k}^{i-j} \right)$$

$$= \frac{x_k^{*2}}{4} A_{i-k}^{i-j} A_{i-k}^{j-k}$$

$$> 0 \qquad \text{(by MBP (4.4))} \tag{D.7}$$

Since $f(s)$ is a cubic polynomial with negative leading coefficient, $f(\tilde{x}_i^{ik}) > 0$ and $f'(\tilde{x}_i^{ik}) > 0$, we can conclude that there is exactly one real root for $s > \tilde{x}_i^{ik}$.

If $A_{i-j}^{i-k} < A_{i-k}^{i-j}$, then from (D.6)

$$f''(\tilde{x}_i^{ik}) = \frac{x_k^*}{2} \left(\frac{(A_{i-j}^{i-k})^2 A_{i-k}^{i-k}}{A_{i-k}^{i-k}} - A_{i-j}^{i-j} A_{i-k}^{i-j} \right)$$

$$= \frac{x_k^*}{2} \left((A_{i-j}^{i-k})^2 - A_{i-j}^{i-j} A_{i-k}^{i-j} \right)$$

$$< \frac{x_k^*}{2} \left((A_{i-k}^{i-j})^2 - A_{i-j}^{i-j} A_{i-k}^{i-j} \right)$$

$$= \frac{x_k^*}{2} A_{i-k}^{i-j} \left(A_{i-k}^{i-j} - A_{i-j}^{i-j} \right)$$

$$= \frac{x_k^*}{2} A_{i-k}^{i-j} A_{j-k}^{i-j}$$

$$< 0 \qquad \text{(by MBP (4.4))} \tag{D.8}$$

Since $f(s)$ is a cubic polynomial with negative leading coefficient, $f(\tilde{x}_i^{ik}) > 0$ and $f''(\tilde{x}_i^{ik}) < 0$, we have $f''(s) < 0$ for $s \geq \tilde{x}_i^{ik}$. This implies that there is exactly one real root for $s > \tilde{x}_i^{ik}$. □

We have the following corollary from (D.7) and (D.8) which will be used later.

Corollary D.2. *For a simple three-strategy coordination game A, either $f'(\tilde{x}_i^{ik}) > 0$ or $f''(\tilde{x}_i^{ik}) < 0$.*

Lemma D.3. *$a_1(s) \geq 0$ for $s \geq \tilde{x}_i^{ik}$. $a_1'(s) \leq 0$ and $a_1''(s) \geq 0$ for $s > \tilde{x}_i^{ik}$.*

Proof of Lemma D.3. Recall

$$a_1(s) = \begin{cases} \frac{1}{36} \frac{(A^{2j-i-k} x(s))^3}{A_{i-k}^{2j-i-k}} & \text{if } x(s) \in \overline{\mathcal{B}}_{jk}^{i} \text{ and } s \geq \tilde{x}_i^{ik} \\ 0 & \text{otherwise} \end{cases}$$

where $x(s) = se_i + (1-s)e_j$.

If $\underline{x}_i^{ij} \leq \tilde{x}_i^{ik}$, then by definition it follows that there does not exist any $s > \tilde{x}_i^{ik}$ with $x(s) \in \overline{\mathcal{B}}_{jk}^{i}$. Therefore in this case we have $a_1(s) = 0$ for $s \geq \tilde{x}_i^{ik}$ and the statement holds vacuously.

Now suppose that $\tilde{x}_i^{ik} < \underline{x}_i^{ij}$. For s such that $\tilde{x}_i^{ik} < s \leq \underline{x}_i^{ij}$, we have $x(s) \in \overline{\mathcal{B}}_{jk}^i$. This implies that $A^{2j-i-k}x(s) \geq 0$. $(x(s))_j < 1 - \tilde{x}_i^{ik} = x_j^*$ and therefore from Lemma 5.3, we have $A_{i-k}^{2j-i-k} > 0$. Differentiating $a_1(s)$ twice for $\tilde{x}_i^{ik} < s < \underline{x}_i^{ij}$, we get

$$
\begin{aligned}
a_1'(s) &= \left(\frac{A_{i-j}^{2j-i-k}}{12}\right) \frac{(A^{2j-i-k}x(s))^2}{A_{i-k}^{2j-i-k}} \\
a_1''(s) &= \frac{(A_{i-j}^{2j-i-k})^2}{6} \frac{A^{2j-i-k}x(s)}{A_{i-k}^{2j-i-k}} \geq 0
\end{aligned}
$$

MBP (4.4) implies that $A_{i-j}^{j-k} < 0 < A_{i-j}^{i-j}$. Since $A_{i-j}^{2j-i-k} = A_{i-j}^{j-k} - A_{i-j}^{i-j}$, it follows that $a_1'(s) \leq 0$. By definition, $A^{2j-i-k}\underline{x}^{ij} = 0$ and for $s > \underline{x}_i^{ij}$, $x(s) \notin \overline{\mathcal{B}}_{jk}^i$. Therefore $a_1(s) = 0$ for $s \geq \underline{x}_i^{ij}$ and from the above arguments it follows that the given statement holds. ◻

Lemma D.4. *$a_2(s) \leq 0$ for $s \geq \tilde{x}_i^{ik}$. $a_2'(s) \leq 0$ and $a_2''(s) \leq 0$ for $s > \tilde{x}_i^{ik}$*

Proof of Lemma D.4. Recall

$$
a_2(s) = \begin{cases} -\frac{1}{36} \frac{(A^{2k-i-j}x(s))^3}{A_{i-j}^{2k-i-j}} & \text{if } x(s) \in \overline{\mathcal{B}}_{kj}^i \text{ and } s \geq \tilde{x}_i^{ik} \\ 0 & \text{otherwise} \end{cases}
$$

where $x(s) = se_i + (1-s)e_j$.

If $\overline{x}_i^{ij} \leq \tilde{x}_i^{ik}$, then by definition it follows that there does not exist any $s > \tilde{x}_i^{ik}$ with $x(s) \in \overline{\mathcal{B}}_{kj}^i$. Therefore in this case we have $a_2(s) = 0$ for $s \geq \tilde{x}_i^{ik}$ and the statement holds vacuously.

Now suppose that $\tilde{x}_i^{ik} < \overline{x}_i^{ij}$. For $\tilde{x}_i^{ik} < s < \overline{x}_i^{ij}$, $x(s) \notin \overline{\mathcal{B}}_{kj}^i$ and therefore by definition $a_2(s) = 0$. For $s \geq \overline{x}_i^{ij}$, $x(s) \in \overline{\mathcal{B}}_{kj}^i$. This implies that $A^{2k-i-j}x(s) \geq 0$ and from Lemma 5.3 (relabeling the indices j and k), we have $A_{i-j}^{2k-i-j} > 0$. Differentiating $a_2(s)$ twice for $s > \overline{x}_i^{ij}$, we get

$$
\begin{aligned}
a_2'(s) &= -\frac{(A^{2k-i-j}x(s))^2}{12} \leq 0 \\
a_2''(s) &= -\left(\frac{A_{i-j}^{2k-i-j}}{6}\right) A^{2k-i-j}x(s) \leq 0
\end{aligned}
$$

By definition, $A^{2k-i-j}\overline{x}^{ij} = 0$ and from the above arguments it follows that the given statement holds. ◻

Using Lemmas D.1, D.3 and D.4 we now prove that the function $g(s)$ has a unique real root for $s > \tilde{x}_i^{ik}$.

From Lemma D.3, we have $a_1(\tilde{x}_i^{ik}) \geq 0$. From Lemma 5.6, it follows that $\tilde{x}^{ik} \notin \overline{\mathcal{B}}_{kj}^{i}$ which implies that $a_2(\tilde{x}_i^{ik}) = 0$. Since $f(\tilde{x}_i^{ik}) > 0$ (from (D.4)), it follows that

$$g(\tilde{x}_i^{ik}) = f(\tilde{x}_i^{ik}) + a_1(\tilde{x}_i^{ik}) > 0. \tag{D.9}$$

Suppose $f'(\tilde{x}_i^{ik}) \leq 0$. From Corollary D.2, we conclude that $f''(\tilde{x}_i^{ik}) < 0$. $f(s)$ is a cubic polynomial with negative leading coefficient. This implies that $f''(s) < 0$ for $s > \tilde{x}_i^{ik}$. Therefore, for $s > \tilde{x}_i^{ik}$, we have $f'(s) < f'(\tilde{x}_i^{ik}) \leq 0$. $g(s) = f(s) + a_1(s) + a_2(s)$, implies $g'(s) < 0$ for $s > \tilde{x}_i^{ik}$ (since $a_1'(s), a_2'(s) \leq 0$). $g'(s) < 0$ for $s > \tilde{x}_i^{ik}$ and $g(\tilde{x}_i^{ik}) > 0$, implies that there is exactly one real root of $g(s) = 0$ for $s > \tilde{x}_i^{ik}$.

Henceforth, we assume that $f'(\tilde{x}_i^{ik}) > 0$. Since $f(s)$ is cubic in s, the leading coefficient of s is negative and $f'(\tilde{x}_i^{ik}) > 0$, we can conclude that there exits a unique real number $s_m > \tilde{x}_i^{ik}$, where $f(s)$ attains local maximum. For $s > s_m$, we have $f'(s) < 0$. Let $\hat{s}$ be the unique value of $s > \tilde{x}_i^{ik}$ which solves $f(s) = 0$. Consider the following exhaustive cases:

Case 1: $\underline{x}_i^{ij} \leq \hat{s} \leq \overline{x}_i^{ij}$: By definition $f(\hat{s}) = 0$. In this case, we have $a_1(\hat{s}) = 0$ and $a_2(\hat{s}) = 0$. From (D.3), we have $g(\hat{s}) = 0$. For $\tilde{x}_i^{ik} < s < \hat{s}$ we have, $f(s) > 0$, $a_1(s) \geq 0$ and $a_2(s) = 0$ (since for such s, $x(s) \notin \overline{\mathcal{B}}_{kj}^{i}$). This implies that $g(s) > 0$ for such values of s. For $s > \hat{s}$ we have, $f(s) < 0$, $a_1(s) = 0$ (since for such s, $x(s) \notin \overline{\mathcal{B}}_{jk}^{i}$) and $a_2(s) \leq 0$. This implies that $g(s) < 0$ for such values of s. It therefore follows that $\hat{s}$ is the unique root of the function $g(s) = 0$ for $s > \tilde{x}_i^{ik}$.

Case 2: $\hat{s} < \underline{x}_i^{ij}$: For $\tilde{x}_i^{ik} < s \leq \hat{s}$, we have $f(s) \geq 0$, $a_1(s) > 0$ and $a_2(s) = 0$ (since for such s, $x(s) \notin \overline{\mathcal{B}}_{kj}^{i}$). This implies that $g(s) > 0$ for such values of s. Since $f'(s) < 0$ for $s \geq \hat{s}$ and $a_1'(s), a_2'(s) \leq 0$, we have $g'(s) < 0$ for $s \geq \hat{s}$. Therefore, there exits a unique real $s^* > \hat{s} > \tilde{x}_i^{ik}$ for which $g(s^*) = 0$.

Case 3: $\hat{s} > \overline{x}_i^{ij}$: This case has the following two subcases:

Case 3a: $s_m < \overline{x}_i^{ij} < \hat{s}$: For $\tilde{x}_i^{ik} < s \leq \overline{x}_i^{ij}$, $f(s) > 0$, $a_1(s) \geq 0$ and $a_2(s) = 0$ (since for such s, $x(s) \notin \overline{\mathcal{B}}_{kj}^{i}$) which implies $g(s) > 0$ for such s. In particular $g(\overline{x}_i^{ij}) > 0$. Since $f(s)$ is cubic with negative leading coefficient and attains a local minimum at s_m, we have $f'(s) < 0$ for $s > s_m$. In this sub-case we have $\overline{x}_i^{ij} > s_m$ and so it follows that $f'(s) < 0$ for $s \geq \overline{x}_i^{ij}$. $g(s) = f(s) + a_1(s) + a_2(s)$, and $a_1'(s), a_2'(s) \leq 0$ implies $g'(s) < 0$ for $s \geq \overline{x}_i^{ij}$. Also $g(\hat{s}) < 0$, since $a_1(\hat{s}) = 0$ and $a_2(\hat{s}) < 0$. So, there is a unique s^* such that $\overline{x}_i^{ij} < s^* < \hat{s}$ which solves $g(s) = 0$.

Case 3b: $s_m \geq \overline{x}_i^{ij}$: For $s \geq \overline{x}_i^{ij}$, $g(s) = f(s) + a_2(s)$ (since $a_1(s) = 0$ for such s as $x(s) \notin \overline{\mathcal{B}}_{jk}^{i}$). We now show that $f''(\overline{x}_i^{ij}) < 0$. From (D.6), we have

$$\begin{aligned}
f''(\overline{x}_i^{ij}) &= \frac{1}{2}\left(\frac{(A_{i-j}^{i-k})^2}{A_{i-k}^{i-k}} A^{i-k}\overline{x}^{ij} - A_{i-j}^{i-j} A^{i-j}\overline{x}^{ij}\right) \\
&= \frac{1}{2}\left(\frac{(A_{i-j}^{i-k})^2}{A_{i-k}^{i-k}} A^{i-k}\overline{x}^{ij} - A_{i-j}^{i-j}(2A^{i-k}\overline{x}^{ij})\right) && (\text{since } A^{i-k}\overline{x}^{ij} = A^{k-j}\overline{x}^{ij})
\end{aligned}$$

$$
\begin{aligned}
&= \frac{A^{i-k}\overline{x}^{ij}}{2}\left(\frac{(A^{i-k}_{i-j})^2}{A^{i-k}_{i-k}} - 2A^{i-j}_{i-j}\right) \\
&= \frac{A^{i-k}\overline{x}^{ij}}{2}\left(\frac{(A^{i-k}_{i-j})^2 - 2A^{i-j}_{i-j}A^{i-k}_{i-k}}{A^{i-k}_{i-k}}\right) \\
&< 0
\end{aligned}
$$

The last inequality follows from the fact that $A^{i-k}_{i-j} < A^{i-j}_{i-j}$ and $A^{i-k}_{i-j} < A^{i-k}_{i-k}$ as A satisfies MBP (4.4). Since $f(s)$ is a cubic polynomial with negative leading coefficient, we have $f''(s) < 0$ for $s \geq \overline{x}^{ij}_i$. $a_2''(s) \leq 0$, implies $g''(s) < 0$ for $s > \overline{x}^{ij}_i$. We also know that $g(\overline{x}^{ij}_i) > 0$, $g(\hat{s}) < 0$. This implies that there is atleast one real root of $g(s) = 0$ for $s > \overline{x}^{ij}_i$. Since we also have $g''(s) < 0$ for $s > \overline{x}^{ij}_i$, we conclude that there is a unique s^* such that $\overline{x}^{ij}_i < s^* < \hat{s}$ which solves $g(s) = 0$. From this we can conclude that there is a unique real $s > \tilde{x}^{ik}_i$ which solves $g(s) = 0$.

Interchanging the roles of indices j and k, we can show that there is a unique state $z^{ik} \in l^{ik}$ with $z^{ik}_i > \tilde{x}^{ij}_i$ such that $W^k(z^{ik}) = W^j(z^{ik})$.

E. Proof of Theorem 6.1

For $x \in \underline{\mathcal{B}}^i_{jk} \cup \underline{\mathcal{B}}^i_{kj}$, we have

$$
W^k(x) = \frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} \quad \text{and} \quad W^j(x) = \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}}
$$

Suppose $\hat{x}^i$ is in $\underline{\mathcal{B}}^i_{jk}$ and on the line l^{ij}. In this case, for x on the line l^{ij} with $x_i \geq \hat{x}_i$ we have $W^k(x) - W^j(x) \leq 0$. Therefore, for such x we have $\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} \leq \frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}}$. Rewriting this equation, we get

$$
A^{i-j}x \geq \left(\frac{A^{i-j}_{i-j}}{A^{i-k}_{i-k}}\right)^{\frac{1}{3}} A^{i-k}x \tag{E.1}
$$

Recall $H(x,u) = L(x,u) + DW^k(x)u$. For $x \in \underline{\mathcal{B}}^i_{jk}$, $L(x,u) = \begin{bmatrix}0 & \frac{1}{4}(A^{i-j}x)^2 & \frac{1}{4}(A^{i-k}x)^2\end{bmatrix}[u]_+$.

We need to show for the states x with $\hat{x}^i_i < x_i < x^{ijk}_i$, that $H(x, e_k - e_i) = 0$ and in the remaining five basic directions that $H(x,u) \geq 0$.

$$
\begin{aligned}
H(x, e_k - e_i) &= \frac{1}{4}(A^{i-k}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{k-i} \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
H(x, e_k - e_j) &= \frac{1}{4}(A^{i-k}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}} A^{i-j}_{k-j} \\
&= \frac{1}{4}(A^{i-k}x)^2 \left(\frac{A^{i-k}_{i-j}}{A^{i-k}_{i-k}} \right) \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))} \\
H(x, e_i - e_j) &= 0 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}} A^{i-k}_{i-j} \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))} \\
H(x, e_i - e_k) &= 0 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}} A^{i-k}_{i-k} \\
&\geq 0 \\
H(x, e_j - e_k) &= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}} A^{i-k}_{j-k} \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))} \\
H(x, e_j - e_i) &= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}} A^{i-k}_{j-i} \\
&\geq \frac{1}{4}\left(\left(\frac{A^{i-j}_{i-j}}{A^{i-k}_{i-k}} \right)^{\frac{1}{3}} (A^{i-k}x) \right)^2 - \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}} A^{i-k}_{i-j} && \text{(from (E.1))} \\
&= \frac{1}{4}(A^{i-k}x)^2 \left(\frac{(A^{i-j}_{i-j})^{\frac{2}{3}} (A^{i-k}_{i-k})^{\frac{1}{3}} - A^{i-k}_{i-j}}{A^{i-k}_{i-k}} \right) \\
&\geq 0
\end{aligned}
$$

The last inequality follows from the fact that $A^{i-j}_{i-j} \geq A^{i-k}_{i-j}$ and $A^{i-k}_{i-k} \geq A^{i-k}_{i-j}$, as A satisfies the MBP (4.4).

F. Proof of Theorem 6.2

We show that at the states x on l^{ij} with $\hat{x}_i \leq x_i < \underline{x}^{ij}_i$ moving in the direction $e_k - e_i$ satisfies the HJB equation. We have for x on the line l^{ij} with $x_i \geq \hat{x}_i$, $W^k(x) - W^j(x) \leq 0$. For x in the region of interest, we therefore have

$$
\frac{1}{12}\frac{(A^{i-k}x)^3}{A^{i-k}_{i-k}} + \frac{1}{36}\frac{(A^{2j-i-k}x)^3}{A^{2j-i-k}_{i-k}} \leq \frac{1}{12}\frac{(A^{i-j}x)^3}{A^{i-j}_{i-j}}
$$

From Lemma 5.3, we have $A^{2j-i-k}_{i-k} \geq 0$. $x \in \overline{\mathcal{B}}^{i}_{jk}$ implies $A^{2j-i-k}x \geq 0$. Therefore, the second term on the left hand side of the above equation is positive. It therefore follows that (E.1) is satisfied in this case as well.

$$L(x,u) = \begin{bmatrix} 0 & \frac{1}{4}(A^{i-j}x)^2 & \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{k-i}x)^2\right) \end{bmatrix} [u]_+$$

$$DW^k(x) = \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k} \tag{F.1}$$

$$\begin{aligned}
H(x, e_k - e_i) &= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{k-i}x)^2\right) + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{k-i} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{k-i} \\
&= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{k-i}x)^2\right) - \frac{(A^{i-k}x)^2}{4} - \frac{(A^{j-k}x - A^{i-j}x)^2}{12} \\
&= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2\right) - \frac{(A^{j-k}x + A^{i-j}x)^2}{12} - \frac{(A^{j-k}x - A^{i-j}x)^2}{12} \\
&= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2\right) - \frac{1}{12}((A^{j-k}x + A^{i-j}x)^2 + (A^{j-k}x - A^{i-j}x)^2) \\
&= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2\right) - \frac{2}{12}\left((A^{i-j}x)^2 + (A^{j-k}x)^2\right) \\
&= 0 \\
H(x, e_i - e_k) &= \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{i-k} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{i-k} \\
&\geq 0 \\
H(x, e_j - e_k) &= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{j-k} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{j-k} \\
&= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{j-k} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}(A^{j-k}_{j-k} + A^{j-i}_{j-k}) \\
&\geq 0 && \text{(by CG (4.3) and MBP (4.4))} \\
H(x, e_j - e_i) &= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{j-i} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{j-i} \\
&= \frac{1}{4}(A^{i-j}x)^2 + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{j-i} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}(A^{j-k}_{j-i} + A^{j-i}_{j-i}) \\
&\geq \frac{1}{4}(A^{i-j}x)^2 - \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{i-j} && \text{(by CG (4.3) and MBP (4.4))}
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{1}{4}\left(\left(\frac{A^{i-j}_{i-j}}{A^{i-k}_{i-k}}\right)^{\frac{1}{3}}(A^{i-k}x)\right)^2 - \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{i-j} \qquad \text{(from (E.1))}\\
&\geq \frac{1}{4}(A^{i-k}x)^2\left(\frac{(A^{i-j}_{i-j})^{\frac{2}{3}}(A^{i-k}_{i-k})^{\frac{1}{3}} - A^{i-k}_{i-j}}{A^{i-k}_{i-k}}\right)\\
&\geq 0
\end{aligned}
$$

$$
\begin{aligned}
H(x, e_k - e_j) &= \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2\right) + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{k-j} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{k-j}\\
&= \frac{1}{4}(A^{i-k}x)^2 + \frac{1}{6}\left((A^{i-j}x)^2 + (A^{j-k}x)^2 - \frac{1}{2}(A^{i-k}x)^2\right) + \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k}_{k-j}\\
&+ \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{k-j}\\
&= \frac{1}{4}(A^{i-k}x)^2\left(1 + \frac{A^{i-k}_{k-j}}{A^{i-k}_{i-k}}\right) + \frac{1}{12}(A^{2j-i-k}x)^2 + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{k-j}\\
&= \frac{1}{4}(A^{i-k}x)^2\frac{A^{i-k}_{i-j}}{A^{i-k}_{i-k}} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{i-j}\\
&= DW^k(x)(e_i - e_j) \qquad \text{(from (F.1))}
\end{aligned}
$$

Since $L(x, e_i - e_j) = 0$, we have $H(x, e_i - e_j) = DW^k(x)(e_i - e_j)$. It therefore follows that, $H(x, e_k - e_j) = H(x, e_i - e_j) = DW^k(x, e_i - e_j)$ for x on l^{ij} with $\hat{x}_i \leq x_i < \underline{x}^{ij}_i$.

G. Proof of Lemma 6.5

From (6.5), we have

$$
c(\underline{x}^{ij}_i) = \frac{1}{4}\frac{(A^{i-k}\underline{x}^{ij})^2}{A^{i-k}_{i-k}}A^{i-k}_{i-j} + \frac{1}{12}\frac{(A^{2j-i-k}\underline{x}^{ij})^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{i-j}
$$

The second term on the right hand side of the above equation is 0 as $A^{2j-i-k}\underline{x}^{ij} = 0$ by definition and the first term is positive from CG (4.3) and MBP (4.4). It follows that $c(\underline{x}^{ij}_i) > 0$. Since $c'(s) > 0$ for $z^{ij}_i \leq s \leq \underline{x}^{ij}_i$, (from (6.6)) $c(z^{ij}_i) < 0$ and $c(\underline{x}^{ij}_i) > 0$, we conclude that there is a unique state $\overline{x}^i$ on l^{ij} with $z^{ij}_i < \overline{x}^i_i < \underline{x}^{ij}_i$ such that $c(\overline{x}^i_i) = 0$ i.e., $DW^k(\overline{x}^i)(e_i - e_j) = 0$.

We now show that there is a unique state $\underline{x}^i$ on l^{ij} with $\tilde{x}^{ik}_i < \underline{x}^i_i < z^{ij}_i$ such that $W^j(\underline{x}^i) = W^k(\overline{x}^i)$. To prove this, we define a function $h(s)$ as follows

$$h(s) = W^j(x(s)) - W^k(\overline{x}^i) \text{ for } x(s) = se_i + (1-s)e_j,\ \tilde{x}_i^{ik} \le s \le z_i^{ij}$$

We have $DW^j(x) = \frac{1}{4}\frac{(A^{i-j}x)^2}{A_{i-j}^{i-j}}A^{i-j}$.

Differentiating $h(s)$ using the above fact, we have for $\tilde{x}_i^{ik} \le s \le z_i^{ij}$, $h'(s) = \frac{1}{4}(A^{i-j}x)^2 > 0$.

$$\begin{aligned} h(z_i^{ij}) &= W^j(z^{ij}) - W^k(\overline{x}^i) \\ &> W^j(z^{ij}) - W^k(z^{ij}) \qquad (DW^k(x)(e_i - e_j) < 0 \text{ for } z_i^{ij} \le x_i < \overline{x}_i^i \text{ implies } W^k(z^{ij}) > W^k(\overline{x}^i)) \\ &= 0 \end{aligned}$$

We next show that $h(\tilde{x}_i^{ik}) < 0$. In order to prove this we make use of the following fact

(G.1) $\qquad A^{i-k}\overline{x}^i > A^{i-k}\tilde{x}^{ik}$

This follows from MBP (4.3). To see this, define a function $p(s) = A^{i-k}x(s)$ where $x(s) = se_i + (1-s)e_j$ for $\tilde{x}_i^{ik} \le s \le \overline{x}_i^i$. A direct computation gives us $p'(s) = A_{i-j}^{i-k}$ which is positive from MBP (4.3). Since $A^{i-k}\overline{x}^i = p(\overline{x}_i^i)$, $A^{i-k}(\tilde{x}^{ik}) = p(\tilde{x}_i^{ik})$, $\overline{x}_i^i > \tilde{x}_i^{ik}$ and $p'(s) > 0$, (G.1) follows.

$$\begin{aligned} h(\tilde{x}_i^{ik}) &= W^j(\tilde{x}^{ik}) - W^k(\overline{x}^i) \\ &= W^j(\tilde{x}^{ik}) - \frac{1}{12}\frac{(A^{i-k}\overline{x}^i)^3}{A_{i-k}^{i-k}} - \frac{1}{36}\frac{(A^{2j-i-k}\overline{x}^i)^3}{A_{i-k}^{2j-i-k}} \\ &\le W^j(\tilde{x}^{ik}) - \frac{1}{12}\frac{(A^{i-k}\overline{x}^i)^3}{A_{i-k}^{i-k}} \qquad (\text{since } A^{2j-i-k}\overline{x}^i \ge 0 \text{ and } A_{i-k}^{2j-i-k} > 0) \\ &\le W^j(\tilde{x}^{ik}) - \frac{1}{12}\frac{(A^{i-k}\tilde{x}^{ik})^3}{A_{i-k}^{i-k}} \qquad (\text{from (G.1)}) \\ &= \frac{1}{12}\frac{(A^{i-j}\tilde{x}^{ik})^3}{A_{i-j}^{i-j}} - \frac{1}{12}\frac{(A^{i-k}\tilde{x}^{ik})^3}{A_{i-k}^{i-k}} \\ &< 0 \qquad (\text{from (D.4)}) \end{aligned}$$

Since $h'(s) > 0$ for $\tilde{x}_i^{ik} \le s \le z_i^{ij}$, $h(\tilde{x}_i^{ik}) < 0$ and $h(z_i^{ij}) > 0$, we conclude that there is a unique state $\underline{x}^i$ on l^{ij} such that $W^j(\underline{x}^i) = W^k(\overline{x}^i)$.

H. Proof of Theorem 6.7

The total cost of the path in (a) is $V(x) = W^k(\overline{x}^i)$. We can find $d(x) > 0$ such that $\overline{x}^i = x + d(x)(e_i - e_j)$. Therefore,

$$V(x) = W^k(\overline{x}^i) = W^k(x + (e_i - e_j)d(x))$$

Since $DW^k(\overline{x}^i)(e_i - e_j) = 0$ (by definition), we have[6]

$$\begin{aligned} DV(x) &= DW^k(\overline{x}^i)(I + (e_i - e_j)Dd(x)) \\ &= DW^k(\overline{x}^i)I + DW^k(\overline{x}^i)(e_i - e_j)Dd(x) \\ &= DW^k(\overline{x}^i) \end{aligned}$$

Recall $H(x,u) = L(x,u) + DV(x)u$. For x in the region of interest

$$L(x,u) = \left[0 \quad \frac{1}{4}(A^{i-j}x)^2 \quad \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2)\right][u]_+ \tag{H.1}$$

We need to show that $H(x, e_i - e_j) = 0$ and for the remaining five basic directions that $H(x,u) \geq 0$.

$$\begin{aligned} H(x, e_i - e_j) &= 0 + DW^k(\overline{x}^i)(e_i - e_j) = 0 \\ H(x, e_j - e_i) &= \frac{1}{4}(A^{i-j}x)^2 + DW^k(\overline{x}^i)(e_j - e_i) \\ &= \frac{1}{4}(A^{i-j}x)^2 + 0 \\ &\geq 0 \\ DW^k(x) &= \frac{1}{4}\frac{(A^{i-k}x)^2}{A^{i-k}_{i-k}}A^{i-k} + \frac{1}{12}\frac{(A^{2j-i-k}x)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k} \\ H(x, e_i - e_k) &= L(x, e_i - e_k) + DV(x)(e_i - e_k) \\ &= 0 + DW^k(\overline{x}^i)(e_i - e_k) \\ &= \frac{1}{4}(A^{i-k}\overline{x}^i)^2 + \frac{1}{12}(A^{2j-i-k}\overline{x}^i)^2 \\ &\geq 0 \\ H(x, e_j - e_k) &= L(x, e_j - e_k) + DV(x)(e_j - e_k) \\ &= \frac{1}{4}(A^{j-k}x)^2 + \frac{1}{4}\frac{(A^{i-k}\overline{x}^i)^2}{A^{i-k}_{i-k}}A^{i-k}_{j-k} + \frac{1}{12}\frac{(A^{2j-i-k}\overline{x}^i)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{j-k} \\ &\geq 0 \end{aligned}$$

$A^{2j-i-k}_{j-k} = A^{j-i}_{j-k} + A^{j-k}_{j-k}$ and therefore the last inequality follows from CG (4.3) and MBP (4.4).

$$\begin{aligned} H(x, e_k - e_j) &= L(x, e_k - e_j) + DV(x)(e_k - e_j) \\ &= L(x, e_k - e_j) + DV(x)(e_k - e_i) + DV(x)(e_i - e_j) \end{aligned}$$

[6] For $Dd(x)$ to make sense, we need to define $d(x)$ in a neighborhood of aff X. To avoid this technicality we proceed as follows: We denote the line joining x^* and $\overline{x}^i$ by $l_{x^*\overline{x}^i}$. Let x be any point in the affine hull of $\underline{x}^i$, x^* and $\overline{x}^i$. For such an x, there exists a unique $d(x) \geq 0$ such that $x + d(x)(e_i - e_j) \in l_{x^*\overline{x}^i}$. With this definition of $d(x)$ its derivative $Dd(x)$ is defined in a neighborhood.

$$\begin{aligned}
&= L(x, e_k - e_j) + DV(x)(e_k - e_i)\\
&= L(x, e_k - e_i) + DV(x)(e_k - e_i)\\
&= H(x, e_k - e_i)
\end{aligned}$$

The third equality follows from the fact that $DV(x) = DW^k(\overline{x}^i)$ and $DW^k(\overline{x}^i)(e_i - e_j) = 0$. The fourth equality follows from the definition of $L(x,u)$ (see (H.1)).

$$\begin{aligned}
H(x, e_k - e_i) &= \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2) + \frac{1}{4}\frac{(A^{i-k}\overline{x}^i)^2}{A^{i-k}_{i-k}}A^{i-k}_{k-i} + \frac{1}{12}\frac{(A^{2j-i-k}\overline{x}^i)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{k-i}\\
&= \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2) - \frac{1}{4}(A^{i-k}\overline{x}^i)^2 - \frac{1}{12}(A^{j-k}\overline{x}^i - A^{i-j}\overline{x}^i)^2\\
&= \frac{1}{6}(A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 - \frac{1}{12}(3(A^{i-k}\overline{x}^i)^2 + (A^{j-k}\overline{x}^i - A^{i-j}\overline{x}^i)^2)\\
&= \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 - ((A^{i-j}\overline{x}^i)^2 + (A^{j-k}\overline{x}^i)^2 + (A^{i-k}\overline{x}^i)^2))
\end{aligned}$$

The last equality in the above set of equations follows from the following simple fact: For $a, b, c \in \mathbb{R}$, with $a + b + c = 0$, we have

(H.2) $$3c^2 + (b-a)^2 = 2(a^2 + b^2 + c^2)$$

The proof will be complete if we show that for x on l^{ij} with $\underline{x}^i_i < x_i < \overline{x}^i_i$, $H(x, e_k - e_i) \geq 0$ i.e.,

(H.3) $$(A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 \geq (A^{i-j}\overline{x}^i)^2 + (A^{j-k}\overline{x}^i)^2 + (A^{i-k}\overline{x}^i)^2$$

We now prove that (H.3) is satisfied for x in the region of interest. We define a function $f_1(s)$ as follows

(H.4) $$f_1(s) = 3(A^{i-k}x(s))^2 + (A^{j-k}x(s) - A^{i-j}x(s))^2$$

for $x(s) = se_i + (1-s)e_j$ where $s \in \mathbb{R}$.

Using (H.2), it follows that $f_1(s) = 2((A^{i-j}x(s))^2 + (A^{j-k}x(s))^2 + (A^{i-k}x(s))^2)$. Clearly $f_1(s)$ is a quadratic function in s with positive leading coefficient and hence strictly convex. Therefore, to prove (H.3), it is sufficient to show that $f_1'(\overline{x}^i_i) \leq 0$. In what follows, we will prove this.

Differentiating (H.4), we get

$$\begin{aligned}
f_1'(s) &= 6A^{i-k}_{i-j}A^{i-k}x(s) + 2(A^{j-k}_{i-j} - A^{i-j}_{i-j})(A^{j-k}x(s) - A^{i-j}x(s))\\
&= 6A^{i-k}_{i-j}A^{i-k}x(s) - 2A^{i+k-2j}_{i-j}A^{2j-i-k}x(s)
\end{aligned}$$

Note that $A^{i+k-2j}_{i-j} = A^{i-j}_{i-j} + A^{k-j}_{i-j}$ which is positive from CG (4.3) and MBP (4.4).

By definition, $DW^k(\overline{x}^i)(e_i - e_j) = 0$ i.e.,

$$3\frac{A_{i-j}^{i-k}}{A_{i-k}^{i-k}}(A^{i-k}\overline{x}^i)^2 - \frac{A_{i-j}^{i+k-2j}}{A_{i-k}^{2j-i-k}}(A^{2j-i-k}\overline{x}^i)^2 = 0 \tag{H.5}$$

We now compute as follows

$$\begin{aligned}
f_1'(\overline{x}_i^i) &= 6A_{i-j}^{i-k}A^{i-k}\overline{x}^i - 2A_{i-j}^{i+k-2j}A^{2j-i-k}\overline{x}^i \\
&= 6A_{i-j}^{i-k}A^{i-k}\overline{x}^i - 2A_{i-j}^{i+k-2j}\left(\frac{3A_{i-j}^{i-k}A_{i-k}^{2j-i-k}}{A_{i-k}^{i-k}A_{i-j}^{i+k-2j}}\right)^{\frac{1}{2}} A^{i-k}\overline{x}^i && \text{(from (H.5))} \\
&= 2(A^{i-k}\overline{x}^i)\left[3A_{i-j}^{i-k} - \left(\frac{3A_{i-j}^{i-k}A_{i-k}^{2j-i-k}A_{i-j}^{i+k-2j}}{A_{i-k}^{i-k}}\right)^{\frac{1}{2}}\right]
\end{aligned}$$

Since $A^{i-k}\overline{x}^i > 0$, the above expression is non-positive if and only if

$$3A_{i-j}^{i-k} \leq \left(\frac{3A_{i-j}^{i-k}A_{i-k}^{2j-i-k}A_{i-j}^{i+k-2j}}{A_{i-k}^{i-k}}\right)^{\frac{1}{2}}$$

Squaring the above equation and rearranging the terms, we see that the above condition is equivalent to $3A_{i-j}^{i-k}A_{i-k}^{i-k} \leq A_{i-j}^{i+k-2j}A_{i-k}^{2j-i-k}$ which follows directly from the necessary condition (6.9) for the retreating case. This implies that $f_1'(\overline{x}_i^i) \leq 0$ and so the proof is complete.

The argument above in fact proves a stronger statement than (H.3). We present it as a corollary which will be used in the proof of Theorem 6.8.

Corollary H.1. *In the retreating case (under the assumptions in Theorem 6.7) for any two states x and z with $\underline{x}_i^i < x_i \leq z_i < \overline{x}_i^i$ in the region spanned by the points $\underline{x}^i$, x^* and $\overline{x}^i$ such that the line joining x and z is parallel to $e_i - e_j$, we have*

$$(A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 \geq (A^{i-j}z)^2 + (A^{j-k}z)^2 + (A^{i-k}z)^2$$

Proof of Corollary H.1. If x and z are on l^{ij} then we are done from the above analysis. Suppose that x and z are any two states satisfying the assumptions but not on l^{ij}. Extend the line joining the states x^* and x until it intersects the line l^{ij} (call this state $\check{x}$). Similarly, extend the line joining the states x^* and z until it intersects the line l^{ij} (call this state $\check{z}$). Clearly, we have $\check{x}_i \leq \check{z}_i$. By construction and from our given assumptions we can find a real number t with $0 < t < 1$ such that

$$x = (1-t)\check{x} + tx^*$$

$$z = (1-t)\check{z} + tx^*$$

We now have

$$
\begin{aligned}
(A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 &= (1-t)^2\left[(A^{i-j}\check{x})^2 + (A^{j-k}\check{x})^2 + (A^{i-k}\check{x})^2\right] \\
&\geq (1-t)^2\left[(A^{i-j}\check{z})^2 + (A^{j-k}\check{z})^2 + (A^{i-k}\check{z})^2\right] \qquad \text{(since } \check{x}, \check{z} \in l^{ij}\text{)} \\
&= (A^{i-j}z)^2 + (A^{j-k}z)^2 + (A^{i-k}z)^2
\end{aligned}
$$

□

I. Proof of Theorem 6.8

Let x be any state in $(R_2^1)^\circ$. We have $V(x) = W^k(z(x))$ where $z(x) - x = (e_i - e_j)d(x)$ for some $d(x) > 0$. Since $z(x)$ is on the line l^{ik}, we have $d(x) = x_j = e_j'x$. A direct computation gives us

$$
\begin{aligned}
DV(x) &= DW^k(z)D(x + (e_i - e_j)d(x)) \\
&= DW^k(z)(I + (e_i - e_j)Dd(x)) \\
&= DW^k(z)(I + (e_i - e_j)e_j') \qquad \text{(since } d(x) = e_j'x\text{)}
\end{aligned}
$$

From the above equation, it follows that

$$
\begin{aligned}
DV(x)(e_i - e_j) &= 0 \\
DV(x)(e_i - e_k) &= DW^k(z)(e_i - e_k) \\
DV(x)(e_j - e_k) &= DW^k(z)(e_i - e_k)
\end{aligned}
$$

From the above two equations, we get $H(x, e_k - e_i) = H(x, e_k - e_j)$. We now have

$$
\begin{aligned}
H(x, e_i - e_j) &= L(x, e_i - e_j) + DV(x)(e_i - e_j) \\
&= 0 \\
H(x, e_j - e_i) &= L(x, e_j - e_i) + DV(x)(e_j - e_i) \\
&= \frac{1}{4}(A^{i-j}x)^2 + 0 \\
&\geq 0 \\
H(x, e_i - e_k) &= L(x, e_i - e_k) + DV(x)(e_i - e_k) \\
&= 0 + DW^k(z)(e_i - e_k) \\
&= \frac{1}{4}(A^{i-k}z)^2 + \frac{1}{12}(A^{2j-i-k}z)^2 \\
&\geq 0 \\
H(x, e_j - e_k) &= L(x, e_j - e_k) + DV(x)(e_j - e_k) \\
&= \frac{1}{4}(A^{j-k}x)^2 + \frac{1}{4}\frac{(A^{i-k}z)^2}{A^{i-k}_{i-k}}A^{i-k}_{j-k} + \frac{1}{12}\frac{(A^{2j-i-k}z)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{j-k} \\
&\geq 0
\end{aligned}
$$

$$
\begin{aligned}
H(x, e_k - e_i) &= L(x, e_k - e_i) + DV(x)(e_k - e_i) \\
&= \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2) + \frac{1}{4}\frac{(A^{i-k}z)^2}{A^{i-k}_{i-k}}A^{i-k}_{k-i} + \frac{1}{12}\frac{(A^{2j-i-k}z)^2}{A^{2j-i-k}_{i-k}}A^{2j-i-k}_{k-i} \\
&= \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2) - \frac{1}{4}(A^{i-k}z)^2 - \frac{1}{12}(A^{j-k}z - A^{i-j}z)^2 \\
&= \frac{1}{6}(A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 - \frac{1}{12}(3(A^{i-k}z)^2 + (A^{j-k}z - A^{i-j}z)^2) \\
&= \frac{1}{6}((A^{i-j}x)^2 + (A^{j-k}x)^2 + (A^{i-k}x)^2 - ((A^{i-j}z)^2 + (A^{j-k}z)^2 + (A^{i-k}z)^2)) \\
&\geq 0
\end{aligned}
$$

The last inequality follows from Corollary H.1.

References

Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and economic behavior*, 5(3):387–424.

Blume, L. E. (2003). How noise matters. *Games and Economic Behavior*, 44(2):251–271.

Cho, I.-K., Williams, N., and Sargent, T. J. (2002). Escaping nash inflation. *The Review of Economic Studies*, 69(1):1–40.

Dokumacı E. and Sandholm, W. H. (2011). Large deviations and multinomial probit choice. *Journal of Economic Theory*, 146(5):2151 – 2158.

Kandori, M., Mailath, G. J., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, pages 29–56.

Kandori, M. and Rob, R. (1998). Bandwagon effects and long run technology choice. *Games and Economic Behavior*, 22(1):30 – 60.

Myatt, D. P. and Wallace, C. (2003). A multinomial probit model of stochastic evolution. *Journal of Economic Theory*, 113(2):286–301.

Sandholm, W. H. and Staudigl, M. (2016). Large deviations and stochastic stability in the small noise double limit. *Theoretical Economics*, 11(1):279–355.

Sandholm, W. H., Tran, H. V., and Arigapudi, S. (2018). Hamilton-jacobi equations with semilinear costs and state constraints, with applications to large deviations in games.

Williams, N. (2001). *Escape dynamics in learning models*. PhD thesis, University of Chicago, Dept. of Economics.

Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pages 57–84.