Convergence of Learning Algorithms in Auction Markets and Bimatrix Games

Shahram Tabandeh

Department of Electrical & Computer Engineering

McGill University

Montreal, Canada

2014-8-14

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

DEDICATION

To my family.

ACKNOWLEDGEMENTS

I would like to thank Professor Fabrice Labeau for his help during my research at McGill. He made very useful comments that greatly improved my work, and he was always available for discussions. I believe that he is a great scientist and, at the same time, a great manager. I would also like to thank Professor Michalska for the time that I worked with her.

I have much gratitude to Dr. Melita Hadzagic, and Sara Kozlowska who have dedicated many hours to my work, made excellent comments about the text, its structure, and notations.

I am happy that I studied in Montreal. I met a lot of wonderful people including, Adrian Ngoly, Dr. Vladislav Ganine, and Dr. Mu Chiao Lu. I also enjoyed being with wonderful pets, Sputnik and my dear Tyrone.

Finally, I would like to say that no accomplishment would have been possible if it was not for the huge love that I received from my parents Massi and Majid Tabandeh.

ABSTRACT

This thesis addresses calculation of Nash equilibria in a bimatrix game as well as convergence in algorithmic learning for two categories of discrete and discontinuous games, namely bimatrix games and double auction markets.

First, a method which does not use estimation is proposed for calculating a Nash equilibrium of the bimatrix game when the fictitious play converges to a Shapley polygon. By using a lexicographical order, the existence of a Nash equilibrium for every Shapley polygon in a bimatrix game is proven. The calculation of the Nash equilibrium follows from a simple matrix reduction technique which imposes no assumption about the matrices of the bimatrix game.

Second, an evolutionary algorithm is proposed for learning in double auction markets where the buyers and sellers follow the most successful member of their respective population in the previous round of the game and mutate their bids by a diminishing Gaussian distribution. The existence of a sequence of Gaussian distributions, such that the stochastic learning algorithm converges to a Nash equilibrium, is proven for risk neutral and risk averse players.

Finally, an evolutionary random search algorithm is proposed for learning of the optimum bid in double auction markets where the agents are considered as members of either the population of sellers or the population of buyers. Participants attempt to learn the optimum bid prices that maximize their individual gain in the next round of the game. The convergence of the learning algorithm to a Nash equilibrium of the game is analyzed, and performance of the algorithm is compared with the performance of the genetic learning algorithm previously used for the same purpose.

ABRÉGÉ

La thèse aborde le problème relatif au calcul des équilibres de Nash et à celui de la convergence dans l'apprentissage algorithmique de deux catégories de jeux discrets et discontinus: les jeux bimatriciels et les doubles enchères.

Premièrement, la méthode qui n'utilise pas l'estimation, est proposée pour le calcul d'un équilibre de Nash du jeu bimatriciel, quand le jeu fictif converge vers un polygone de Shapley. En utilisant un ordre lexicographique, on a prouvé que dans un jeu bimatriciel pour chaque polygone de Shapley un équilibre de Nash existe. Le calcul de l'équilibre de Nash découle d'une technique simple de réduction de matrice qui n'impose aucune hypothèse sur les matrices de jeu bimatriciel.

Deuxièmement, un algorithme évolutionniste est proposé pour l'apprentissage en cas d'une double enchère où les acheteurs et les vendeurs suivent le membre le plus réussi de leur population respective du tour précédent. Les offres des joueurs sont mutées par une distribution gaussienne qui diminue dans le temps. Pour les joueurs averses au risque et les joueurs neutres, l'existence d'une séquence des distributions gaussiennes, a été prouvée, telles que l'algorithme d'apprentissage stochastique converge vers un équilibre de Nash.

Finalement, un algorithme évolutionniste de recherche aléatoire servant à l'apprentissage de l'offre optimale dans le cas d'une double enchère est proposé. Dans cet algorithme, les agents sont considérés comme membres d'une population de vendeurs ou d'acheteurs. Les participants essaient d'apprendre leurs offres optimales qui maximisent leurs revenus dans le tour de jeu suivant. La convergence de l'algorithme d'apprentissage à un équilibre de Nash du jeu a été analysée et la performance de l'algorithme a été comparée avec la performance de l'algorithme génétique utilisé aux mêmes fins que précédemment.

CLAIMS OF ORIGINALITY

- A practical method is offered to calculate a Nash equilibrium of the game if fictitious play has converged to a Shapley polygon. The existence of an answer is proven by using a lexicographical order.
- An evolutionary algorithm is proposed for learning in double auction markets where buyers and sellers follow the best members of their populations from the previous round of the game and mutate their bids by a diminishing Gaussian distribution. The existence of a sequence of Gaussian distributions, such that the stochastic learning algorithm converges to a Nash equilibrium, is proven for risk neutral and risk averse players. This research is partially published in [1] [2].
- A stochastic learning algorithm is suggested that imitates a trader's behaviour in a double auction market. The algorithm is designed based on the fundamental characters of the utility functions for buyers and sellers. For example, a buyer's maximum utility occurs when the bid is closest to the price that a seller who is matched with this buyer bids to sell. The buyers bid to buy at the minimum price and the sellers offer to sell at the maximum price, but they both try to trade with a specific proportion of their counterparts. The convergence of this behavioural algorithm to a Nash equilibrium is analyzed. This research is partially published in [3].

CONTRIBUTIONS OF CO-AUTHOR

The ideas reflected in the thesis almost entirely belong to the doctoral candidate. He has co-authored papers [1], [2], and [3] with Professor Michalska. She provided advice, added to the explanations, and helped with editing the text in Chapters 4 and 5.

However, proofs of convergence for the algorithms in these two chapters were later added by the candidate.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
Introduction

The goal of game theory is to analyze situations of the so-called *strategic game*, in which one player's best action is dependent on expectations about what other players choose to do [4]. The theory of games of strategy may be described as a mathematical theory of decision making [5–7]. In a strategic game, two or more players play to maximize their outcome known in terms of the value of the *utility function* [8]. In many realistic situations, it is more feasible to learn than to calculate the best player's action, i.e., the one that maximizes its utility function. Players can use their past experience to learn to play better in future rounds only if the game is repeated [9–13]. The past experience used in learning can come from either the player itself or from other players in the game [14].

Learning in strategic games may be achieved using *evolutionary algorithms*, which are population based optimization algorithms inspired by the evolution process as described in biology. An evolutionary algorithm consists of two main processes: one that creates the new candidates to represent the new generation and the other one that decides whether or not to accept them [15–18].

A strategic game may be either cooperative or non-cooperative. As defined by Nash and redefined by Harsanyi [19], a *Cooperative game* is defined as a game where the players are allowed to have binding agreements and binding commitments. In contrast, in a *non-cooperative game*, binding agreements are not allowed. In this thesis, the strategic non-cooperative games are investigated.

Nash equilibrium is a central concept in analyzing non-cooperative games. A Nash equilibrium occurs if none of the players in a game has an incentive to deviate from his/her choice unilaterally [20]. The theory of learning and evolution in games provides an explanation about players' behaviour in disequilibrium and analyzes if an equilibrium is reached as a result of a certain learning or evolution procedure. Moreover, if existent, the time efficient calculation of equilibria is of primordial importance in applied economics [21].

This dissertation is written on the convergence of learning algorithms in auction markets and bimatrix games.

Auction markets are widely used in transaction media. Examples include stock markets, online shopping, and electricity markets. However, the discontinuity of the utility functions of the participants in an auction introduces a significant challenge in designing of the learning algorithms for auction markets.

A *bimatrix game* is a discrete game where the utility function of each players is given by a matrix. It is well known that in a bimatrix game, learning by *fictitious play*, which is a learning rule where a player's action improves towards his/her best action at the current situation, may not converge to an equilibrium of the game but to a cyclic behaviour called a *Shapley polygon*. In the present dissertation, a proof based on lexicographical order is offered which shows that the existence of a Shapley polygon is a sufficient condition for the existence of a *Nash equilibrium*, where no player intends to deviate unilaterally.

Before describing more in detail the problems of interest, the preliminaries are reviewed.

1.1 Preliminaries

A finite game Γ in normal form is defined by the triplet $\Gamma < I, S, U >$ where $I = \{1, ..., p\}$ is the set of players, S is the Cartesian product of the finite sets $S_i = \{s_i^1, ..., s_i^{m_i}\}$ which represents the m_i choices available to player $i \in I$, and U is the set of utility functions $u_i : S \to \Re$ for all $i \in I$.

For player $i \in I$ a mixed strategy x_i is defined as a probability distribution on the set S_i, i.e., $x_i = (x_i^1, x_i^2, ..., x_i^{m_i})$, such that for every $j \in \{1, ..., m_i\}$ $x_i^j \geq 0$ and $\Sigma_{j=1}^{m_i} x_i^j = 1$. Thus, the space of mixed strategies for player i is defined as $\Delta_{x_i} = \{x \in R_+^{m_i} | \Sigma_{j=1}^{m_i} x_i^j = 1\}$ and is visualized as a polytope. A mixed strategy is seen as a randomization among the options available to the player [6, 22]. Action of the player i is referred to as a pure strategy, if he/she chooses an $s_i^j \in S_i$ instead of randomization on the set S_i.

To calculate the expected utility for a mixed strategy, let $x = (x_1, ..., x_p)$ be a profile of the mixed strategies for all players and $s = (\hat{s}_1, ..., \hat{s}_p)$, where $\hat{s}_i \in S_i$, be an arbitrary combination of pure strategies. If players independently randomize on their mixed strategies the probability for this combination of pure strategies is calculated as

$$P(s) = \prod_{j=1}^{n} x^j(\hat{s}_j) \tag{1.1}$$

where $x^j(\hat{s}_j)$ is the probability that player j chooses $\hat{s}_j$, then the expected utility for player i is calculated as

$$u_i(x) = \sum_{s \in S} P(s) u_i(s) \tag{1.2}$$

where S is the set of all possible profiles of pure strategies for all players. Furthermore, the best response function for player $i \in I$ is defined as

$$BR_i(x_{-i}) = argmax_{x_i \in \Delta_{x_i}} u_i(x) \tag{1.3}$$

where x_i is the mixed strategy for player i, Δ_{x_i} is the set of all mixed strategies for player i and x_{-i} signifies the set of mixed strategies of every player but the player i.

Definition 1.1 *A profile of mixed strategies $x = (x_1, ..., x_n)$ is a Nash equilibrium if and only if [20]:*

$$x_i \in BR_i(x_{-i}). \tag{1.4}$$

A bimatrix game is defined by two matrices $A_{n \times m}$ and $B_{n \times m}$ where n is the cardinality for the set of choices for player 1, S_1 and m is the cardinality of the set of choices for player 2, S_2. If player 1 chooses $s_1^i \in S_1$ and player 2 chooses $s_2^j \in S_2$ where $i \in \{1, ..., n\}$ and $j \in \{1, ..., m\}$ then the entry a_{ij} of matrix A is the value of utility for player 1 and the entry b_{ij} of matrix B is the value of utility for player 2.

In this bimatrix game, if player 1, chooses mixed strategy $x_1 \in \Delta_{x_1}$, where $\Delta_{x_1} = \{\zeta \in R_+^n | \Sigma_{i=1}^n \zeta_i = 1\}$ and player 2 chooses $x_2 \in \Delta_{x_2}$, where $\Delta_{x_2} = \{\lambda \in R_+^m | \Sigma_{i=1}^m \lambda_i = 1\}$ then the utilities for both players are calculated as

$$u_1(x_1, x_2) = x_1^T A x_2 \quad \text{and} \quad u_2(x_1, x_2) = x_2^T B x_1. \tag{1.5}$$

The pair (x_1^*, x_2^*) where $x_1^* \in \Delta_{x_1}$ and $x_2^* \in \Delta_{x_2}$ is a Nash equilibrium if and only if for every $x_1 \in \Delta_{x_1}$ and $x_2 \in \Delta_{x_2}$ (see Definition 1.1):

$$x_1^{*T} A x_2^* \geq x_1^T A x_2^* \quad \text{and} \quad x_1^{*T} B x_2^* \geq x_1^{*T} B x_2. \tag{1.6}$$

In economics, utility is a measure of relative satisfaction or desire for consumption, receiving a service or a profit [4]. Utility can be defined as ordinal or cardinal. While ordinal utility explains preference relationships, cardinal utility contains further information about the magnitude of preference among the available choices. Hence, for player i, utility function $u_i : S \to R$ is defined as an order-homomorphism that preserves an order over the set of possible choices S_i, [4, 23, 24].

Based on these assumptions on utility function, a rational choice is defined as a choice which is at least as good as every other choice according to the decision maker's preferences given his/her information [4, 25].

Definition 1.2 *A learning rule is defined as a function that evolves the mixed strategy of every player i, i.e., $L_i : x(k) \to x_i(k+1)$ where $k \in Z^+$ is the discrete time, $x(k)$ is the set of mixed strategies of all players in time k and $x_i(k+1)$ is the mixed strategy of player i in time $k+1$. Similarly, for the continuous time, $t \in R^+$, a learning rule can be defined as $L_i : x(t) \to \dot{x}_i(t)$, where $x(t)$ is the set of mixed strategies of all players in time t and $\dot{x}_i(t)$ is the first derivative of the function $x(t)$, mixed strategy of player i, with respect to time. A learning rule is meaningful if player i, given his/her information, expects that by this change he/she will gain more utility in a specified future.*

1.2 Issues in Learning in Games

This section reviews the most important issues in learning, evolution, and computation of Nash equilibria.

Convergence to Equilibria: it is of primordial importance to know whether the repeated game will converge to a Nash equilibrium, because:

i) Some of the learning rules such as fictitious play which were originally designed to calculate the Nash equilibria of the games [26] may not converge to a Nash equilibrium.

ii) An analysis of players' behaviour when the game is out of equilibrium may be necessary. In particular, the desired is to know what occurs if players choose a specific learning or an evolution rule.

Selection of Equilibria: Two different learning rules may converge to different Nash equilibria of the game even if the learning process starts from the same initial profile of mixed strategies. The selection of equilibria explains which equilibrium will be chosen as a result of a specified learning process. From computational point of view it is important to know if all the Nash equilibria of a game can be calculated by a specific learning rule and from a player's point of view, selection of equilibrium will result in different payoffs for the player [27].

Complexity of Computation: Computation of equilibria is a central question in applied economics that provides answers about price formation in an auction market. Calculation of equilibria often involves solving complex problems. Even in a bimatrix game the degree of complexity of computation increases exponentially with the dimensions of matrices [28].

Sensitivity Analysis of the Algorithms: Many algorithms have a set of parameters to be chosen by the user at their initialization. Thus, it is important to know how a different choice for these parameters changes an algorithm's behaviour in terms of convergence to equilibria, speed of convergence, probability of selection of the different equilibria, etc.

1.3 Thesis Outline

In *Chapter 2*, the previous works related to the problems of interest in this thesis are reviewed.

In *Chapter 3*, fictitious play [26, 29] may not converge to a Nash equilibrium but a Shapley polygon [30]. The existence of a Nash equilibrium for every Shapley polygon in a bimatrix game is proven and a simple matrix reduction technique for its calculation is proposed.

In *Chapter 4*, an evolutionary algorithm is proposed for learning in double auction markets where the buyers and sellers follow the best members of their populations from the previous round of the game and mutate their bids by a diminishing Gaussian distribution.

In *Chapter 5*, a novel stochastic algorithm which mimics the behaviour of buyers and sellers in a real market is suggested for learning in double auction markets. The convergence of the proposed algorithm under certain conditions is proven.

In *Chapter 6*, concluding remarks are presented and possibilities for continuation of this research are reviewed.

CHAPTER 2
Literature Survey

As indicated by Charles Darwin in "the Origin of Species by Means of Natural Selection" [31], natural selection gives more chance to the most fit members of the population to survive and reproduce. Hence, nature has the power of selection as humans do. Darwin's theory states that major cases of variability are recognized as follows [31]: the effect of habit, e.g., climate; abundance of food; use or disuse of parts; and the correlated variation which follows from the crossing of different species.

In "The Selfish Gene" [32], Richard Dawkins promotes the idea that adaptation and natural selection is a gene-centred evolution pursued by selfish and utility-oriented genes. In contrast, group selection theory is based on a group's benefit [33]. This dissertation follows the first idea. Evolutionary game theory has its roots in biology [15, 34–37].

A player may use an optimization technique when learning in a game, because he/she wishes to improve his/her strategy. Moreover, there are learning techniques which are specifically designed for learning and evolution in games. A list of the most used learning and evolutionary rules is presented in Section 2.1.

2.1 Learning and Evolution in Games

In this Subsection, famous rules of learning in games [38–42] such as best replay dynamics [14], fictitious play [26, 29], stochastic fictitious play [43], and reinforcement learning [44] are introduced. All learning rules are presented for a bimatrix game defined by matrices $A_{n\times m}$ and $B_{n\times m}$.

Partial Best Response: A discrete time partial best response dynamics is defined by the learning rule [14, 35]:

$$x_1(k+1) = \lambda BR_1(x_2(k)) + (1-\lambda)x_1(k) \quad \text{and} \quad x_2(k+1) = \lambda BR_2(x_1(k)) + (1-\lambda)x_2(k) \quad (2.1)$$

where $BR_1(.)$ and $BR_2(.)$ are the best response functions for the players 1 and 2, λ is the fraction of population that implements the change, $k \in \Re^+$ is the discrete time, $x_1(0) \in \Delta_{x_1}$ and $x_2(0) \in \Delta_{x_2}$ are the initial mixed strategies, and Δ_{x_1}, Δ_{x_2} are the mixed strategy polytopes for both players. Similarly, continuous time partial best response is defined by

$$\dot{x}_1(t) = \lambda(BR_1(x_2(t)) - x_1(t)) \quad \text{and} \quad \dot{x}_2(t) = \lambda(BR_2(x_1(t)) - x_2(t)) \tag{2.2}$$

where $t \in \Re^+$ is the continuous time and $(x_1(0), x_2(0)) \in \Delta_{x_1} \times \Delta_{x_2}$.

Fictitious Play: The fictitious play as introduced by [26,29,45–47] is the oldest and best known learning rule in games. In fictitious play, each player chooses his/her best response according to his/her observation of the historic actions of other players.

A simple formulation of what happens in the fictitious play is given by

$$x_1(k+1) = \frac{BR_1(x_2(k)) + kx_1(k)}{k+1} \quad \text{and} \quad x_2(k+1) = \frac{BR_2(x_1(k)) + kx_2(k)}{k+1} \tag{2.3}$$

where $BR_1(.)$ and $BR_2(.)$ are the best response functions for the players 1 and 2, $k \in \Re^+$, is the discrete time, $x_1(0) \in \Delta_{x_1}$ and $x_2(0) \in \Delta_{x_2}$ are the initial mixed strategies, Δ_{x_1} and Δ_{x_2} are the mixed strategy polytopes for both players. Similarly, the continuous time fictitious play is defined as

$$\dot{x}_1(t) = \frac{BR_1(x_2(t)) - x_1(t)}{t} \quad \text{and} \quad \dot{x}_2(t) = \frac{BR_2(x_1(t)) - x_2(t)}{t} \tag{2.4}$$

where $t \in \Re^+$ is the continuous time and $(x_1(0), x_2(0)) \in \Delta_{x_1} \times \Delta_{x_2}$.

Stochastic Fictitious Play: In stochastic fictitious play [43] the utility function of players is manipulated by Harsanyi's random shocks [48], which means that instead of the game $\Gamma < I, S, U >$ the perturbed game $\Gamma^* < I, S, U >$ is considered for the learning process. It is proven that convergence of both beliefs and strategies happens for a wider range of games than what is observed for the standard fictitious play [49].

Replicator Dynamics: In the replicator dynamics learning rule [14, 50], the share of the players using a strategy grows proportionally to the payoff of that strategy. One of the common approaches is to let population of the players who choose a strategy grow proportional to the logarithm of the value of utility function obtained by players who have chosen this strategy. For a bimatrix game defined by $A_{n\times m}$ and $B_{n\times m}$, the replicator dynamics formulas for x_1 and x_2, mixed strategies of both players are calculated as

$$\dot{x}_1^i = x_1^i((Ax_2)^i - x_1^T Ax_2) \quad \text{and} \quad \dot{x}_2^j = x_2^j((Bx_1)^j - x_2^T Bx_1) \tag{2.5}$$

where $i \in \{1,..,n\}$, $j \in \{1,...,m\}$, and $(Ax_2)^i$, $(Bx_1)^j$ are the ith and the jth entries of the vectors Ax_2 and Bx_1.

Reinforcement Learning: Reinforcement learning is a learning rule usually used by players in uncertain environments [44, 51–54]. There are several components of the model including: i) the policy which is a mapping from the state of the environment to the space of actions; ii) the reward function which is a mapping from the pairs of state-action to the set of real numbers that indicates players' desire; iii) the value function which is a summation over the values of the reward function starting from the given time; iv) a model of the environment which is omitted from many examples of reinforcement learning.

Q-learning as one of the approaches to reinforcement learning is used often for learning in games. Q-learning in a simple form is represented by the equation

$$\Delta Q(s_t, a_t) = \alpha(r + \gamma max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s, a)) \tag{2.6}$$

where $Q(s_t, a_t)$ is a function that values action a_t at the state s_t for time t, α is the learning rate and γ is the discount parameter for future reward. If γ is considered equal to zero, Q-learning will just be a combination of a dynamic programming and a Robinsson-Munro stochastic process.

Bayesian Learning: Bayesian Learning is one the most important tools for statistical machine learning. Players use the Bayes' rule to learn from each others' actions and act upon it to maximize their payoff [11].

2.2 Calculation of Equilibria

Besides learning and evolutionary methods which are used for the calculation of Nash equilibria, there exists another class of algorithms which are also used for the computation of the equilibria [21, 29, 55–57]. These algorithms usually use the general concepts of optimization and mathematical programming to find the Nash equilibria. The most well-known algorithms of this group are:

Lemke-Howson Algorithm: In [55], an algebraic proof is given for the existence of equilibria in bimatrix games. Since the proof is constructive, the algorithm is used to calculate at least one Nash equilibrium of the game. The bimatrix game defined by the matrices A and B is converted to a Linear Complementarity Problem (LCP) which thereafter is solved not on the strategy space but on a nonnegative orthant [58]. The Lemke-Howson algorithm is seen as the classical method for finding a Nash equilibrium of the bimatrix game [55,58], but may never reach some of the equilibria and is recognized as a high complexity algorithm [28].

Van den Elzen-Talman Algorithm: In this algorithm, the problem of finding a Nash equilibrium in a finite game is solved as a stationary point problem [58]. The algorithm can start from almost any mixed strategy. The Lemke-Howson algorithm and the Van den Elzen-Talman algorithm may reach different Nash equilibria even if they start from the same point [59]. Reference [59] also proves that the Van den Elzen-Talman algorithm is a special case of the global Newton method for finding the Nash equilibria [60].

Global Newton method: In this algorithm Kohlbergs structure theory and homotopy are used to find the Nash equilibria through the topological property of the graph of Nash equilibrium. The algorithm works for finite games and it is proven that both Lemke-Howson Algorithm and Van den Elzen-Talman Algorithm are special cases of this algorithm [59].

Algebraic approach to compute the Nash equilibria: Gröbner Bases are used to solve systems of polynomial equations which result in the set of the Nash equilibria of the game, [21]. Since Gröbner Bases are used for the computation of Nash equilibria a lot of redundant polynomials are generated. The positive side of this type of algorithm is that the set of all the equilibria will be found.

Before a literature survey about learning in auctions, different types of auction markets are explained.

2.3 Types of Auction Markets

Auctions have been used since antiquity for selling objects [61–64]. There is historic evidence for auctioning as early as 25 centuries ago. Now, a wide variety of goods and services are sold by different auction types. Goods and services traded through auctioning include government issued bonds and securities [65], the right to use natural resources [66, 67], stock market shares [68], and retail or whole online businesses [69, 70]. References [71–78] describe auction mechanisms. Here, common types of auction markets are described [61, 79–81]:

- English auction or open ascending auction, is perhaps the most known type of auctioning, where there is an object to sell, bidding starts from a low price and it goes up as long as there is more than one buyer interested in buying the object.
- Dutch auction, is an auction where auctioneer starts with a high price and lowers the price until the object is sold.
- Sealed-bid first price auction, is a type of auctioning where the buyers send their bids in closed envelops and the bidder whose bid is the highest gets the object and pays as much as his/her bid.
- Sealed-bid second price auction, is similar to the previous auction but the winner pays as much as the second highest bid.
- Double auction markets are the basic framework of exchange in many markets including the New York Stock Exchange. Double auctions are also used as mechanisms of exchange in electricity markets [82, 83]. In a double auction market, buyers bid and

sellers offer simultaneously. A transaction is possible when a buyer is matched with a seller that is offering a price lower than what the buyer bids to pay. In [84], behaviour of the players in an oral double auction market is studied.

Any of these types can be extended to auctioning for multiple objects or the case in which multiple sellers compete with each other to sell to a single buyer. Each buyer has a value attributed to the object as the maximum price that he/she is willing to pay for and each seller can have a floor for selling the object.

Prior to the problem of the convergence to Nash equilibria, the existence of the equilibria must be addressed. Auctions are examples of discontinuous games [85]. References [86–89] present conditions for the existence of equilibria in discontinuous games. Papers [85, 90] deliver conditions of existence of equilibria in auction markets.

2.4 Learning in Auction Markets

Before introducing the novel ideas for evolution in the next chapters, a list of examples of usual methods for learning in auctions is presented:

1) In [91, 92], learning in a double auction market is studied. A genetic algorithm is used as the learning tool. The convergence of the algorithm is proved and confirmed by simulation results.

2) Reference [93] shows that for first-price auctions, learning by fictitious play converges to an equilibrium if the process is applied for a sufficient long time.

3) In [94, 95] Simulated Annealing and Genetic Algorithm are used for learning in auctions.

4) Since decentralization is important for the electricity markets, multiple references investigate how players can bid better by learning in electricity markets. For example, in [96] an algorithm that uses the minimum information is used for learning.

5) Reference [97] uses the concepts of mean field equilibria for learning in a sequence of second price auctions.

6) Reinforcement learning has also been used for learning in auctions. For example, in [98] Q-learning is used for learning in auctioning of Cognitive Radio Networks.

7) In [99], Bayesian learning is used for learning in an auction, again for Cognitive Radio Networks.

8) In [100], an example of the Tabu Search algorithm is used for learning in a practical auction for a maintenance scheduling problem.

9) Ant Colony heuristic optimization is used for the NP-complete problem of winner determination in combinatorial auctions [101].

2.5 Convergence to a Shapley Polygon in Bimatrix Games

Fictitious play was introduced as a means for calculating Nash equilibria of the game [26, 29]. However, reference [47] shows that the original fictitious play was different than the formulation used by most researchers for learning in games.

In [102], a proof that fictitious play converges to a Nash equilibrium of the game in 2×2 bimatrix games is presented. However, in [103] it is shown that an extra condition on the tie breaking rule is needed to guarantee the convergence. This result was extended to noisy 2×2 games by [49]. In paper [104] convergence of fictitious play to a Nash equilibrium for every $n \times 2$ game is proven. Convergence of fictitious play can not be guaranteed for $n \times m$ games and examples of non convergence are well known [30, 105, 106]. In addition, [107, 108] explain rate of convergence of fictitious play to a Shapley polygon.

In [109], sufficient conditions for convergence of stochastic fictitious play are described. Stochastic approximation is the traditional tool for analyzing convergence of stochastic fictitious play [110, 111]. Reference [105], presents a case of bimatrix games where stochastic fictitious play does not converge to a mixed Nash equilibrium.

Furthermore, papers [50, 112] show that for a 3×3 game of Rock, Paper, Scissor, learning by the rule of replicator dynamics will not result in convergence to the mixed Nash equilibrium of the game.

CHAPTER 3
Existence and Calculation of a Nash Equilibrium for Every Shapley Polygon Resulted by Fictitious Play in Bimatrix Games

3.1 Introduction

Fictitious play was originally proposed for the calculation of Nash equilibria in a bimatrix game [26, 29]. It was then discovered that learning by fictitious play may converge to a Shapley polygon and not to a Nash equilibrium of the bimatrix game [30]. The question is, if fictitious play converges to a Shapley polygon, how is a Nash equilibrium calculated?

Benaïm, Hofbauer, and Hopkins [105], present a game in which the utility matrix is given by

$$A = \begin{bmatrix} 0 & -3 & 1 \\ 1 & 0 & -2 \\ -3 & 1 & 0 \end{bmatrix} \tag{3.1}$$

In this example, none of the usual learning methods, namely, fictitious play, stochastic fictitious play nor reinforcement learning converge to the unique Nash equilibrium of the game, $[9/32, 10/32, 13/32]$. A method called the Time Average of the Shapley Polygon (TASP) is introduced that calculates an estimate of a Nash equilibrium of the bimatrix game if the fictitious play has converged to a Shapley polygon. However, this estimate can be "quite distinct" from any Nash equilibrium of the game [105].

In this chapter, firstly, it is proven that in a bimatrix game, the existence of a Shapley polygon is a sufficient condition for the existence of a Nash equilibrium. Secondly, if fictitious play converges to a Shapley polygon then this fact can be used to calculate a Nash equilibrium of the game.

3.2 Fictitious Play and Shapley Polygons

Let matrices $A_{n\times m}$ and $B_{m\times n}$ define a bimatrix game. If player 1 chooses mixed strategy $x_1 \in \Delta_{x_1}$, and if player 2 chooses mixed strategy $x_2 \in \Delta_{x_2}$, where Δ_{x_1} and Δ_{x_2} are the sets of all the possible mixed strategies for players 1 and 2 respectively, then the utility functions for both players are

$$u_1(x_1, x_2) = x_1^T A x_2 \quad \text{and} \quad u_2(x_1, x_2) = x_2^T B x_1. \tag{3.2}$$

The pair $(\tilde{x}_1, \tilde{x}_2)$ where $\tilde{x}_1 \in \Delta_{x_1}$ and $\tilde{x}_2 \in \Delta_{x_2}$, is a Nash equilibrium if and only if for every $x_1 \in \Delta_{x1}$ and $x_2 \in \Delta_{x_2}$,

$$\tilde{x}_1^T A \tilde{x}_2 \geq x_1^T A \tilde{x}_2 \quad \text{and} \quad \tilde{x}_2^T B \tilde{x}_1 \geq x_2^T B \tilde{x}_1. \tag{3.3}$$

Definition 3.1 *The i-th best response (reply) region [59], $\Delta_{x_1}^i$, $i \in \{1, ..., m\}$, for the set of mixed strategies for player 1, Δ_{x_1}, is defined by $\Delta_{x_1}^i \triangleq \{x \in \Delta_{x_1} | (Bx_1)^i \geq (Bx_1)^{i'}\}$, where $i' \in \{1, ..., m\}$, and $(Bx_1)^i$ denotes the i-th component of the vector Bx_1. Similarly, for $i \in \{1, ..., n\}$, the i-th best response region $\Delta_{x_2}^i$, for the set of mixed strategies for player 2, Δ_{x_2}, is defined by $\Delta_{x_2}^i \triangleq \{x \in \Delta_{x_2} | (Ax_2)^i \geq (Ax_2)^{i'}\}$, where $i' \in \{1, ..., n\}$ and $(Ax_2)^i$ denotes the i-th component of the vector Ax_2.*

Definition 3.2 *(Best responses as attractors): if $x_2(t) \in \Delta_{x_2}^k$, then the unity vector, $e_{1k} \in \Delta_{x_1}$, defined by its l-th component as*

$$e_{1k}^l \triangleq \begin{cases} 1 & l = k \\ 0 & elsewhere \end{cases} \tag{3.4}$$

is a best response or an attractor for player 1 to player 2, i.e., there exists no $x_1 \in \Delta_{x_1}$ such that $x_1^T A x_2(t) > e_{1k}^T A x_2(t)$. Hereafter, the set of all unity vectors $e_{1k} \in \Delta_{x_1}$ is denoted as E_1, where every $e_{1k} \in \Delta_{x_1}$ is a pure strategy for player 1.

If $x_1(t) \in \Delta_{x_1}^{k'}$*, then the unity vector,* $e_{2k'} \in \Delta_{x_2}$*, defined by its l-th component as*

$$e_{2k'}^{l} \triangleq \begin{cases} 1 & l = k' \\ 0 & elsewhere \end{cases} \tag{3.5}$$

is a best response for player 2 to player 1, i.e., there exists no $x_2 \in \Delta_{x_2}$ *such that* $x_2^T Bx_1(t) > e_{2k'}^T Bx_1(t)$*. Hereafter, the set of all unity vectors* $e_{2k'} \in \Delta_{x_2}$ *is denoted as* E_2*, where every* $e_{2k'} \in \Delta_{x_2}$ *is a pure strategy for player 2.*

Proposition 3.1 *[105] (Attractors and piecewise linear trajectories): if for every* $t \in [t_0, t_1]$ *there exist best response regions* $\Delta_{x_1}^{k'} \subset \Delta_{x_1}$ *and* $\Delta_{x_2}^{k} \subset \Delta_{x_2}$ *such that* $x_1(t) \in \Delta_{x_1}^{k'}$ *and* $x_2(t) \in \Delta_{x_2}^{k}$ *then learning by the rule of fictitious play in continuous time, defined by*

$$\dot{x}_1(t) = \frac{e_{1k} - x_1(t)}{t} \quad and \quad \dot{x}_2(t) = \frac{e_{2k'} - x_2(t)}{t} \tag{3.6}$$

forms the following piecewise linear trajectories: $\forall t \in [0, 1]$*,*

$$x_1(t) = \frac{(t - t_0)}{t} e_{1k} + \frac{t_0}{t} x_1(t_0) \quad and \quad x_2(t) = \frac{(t - t_0)}{t} e_{2k'} + \frac{t_0}{t} x_2(t_0) \tag{3.7}$$

where $t \in R^+$ *is the continuous time and* e_{1k}*,* $e_{2k'}$ *are defined in Definition 3.2.*

Proof. The proof of Proposition 3.1 follows from solving the ordinary differential equation (3.6).

Definition 3.3 *A Shapley polygon, [30, 105], with the set of vertices* $V = \{(a_l, b_l) | a_l \in \Delta_{x_1}, b_l \in \Delta_{x_2}\}_{l=1}^{L}$ *and their corresponding attractors* $S = \{(s_{1l}, s_{2l}) | s_{1l} \in E_1, s_{2l} \in E_2\}_{l=1}^{L}$*, where* s_{1l} *is a best response to* b_l *and* s_{2l} *is a response to* a_l*, is formed if and only if the following conditions are met*

i) For every vertex $(a_l, b_l) \in V$*, where* $l \in \{1, ..., L-1\}$*, there exist* $t > t_0$*,* $k \in \{1, ..., n\}$*,* $k' \in \{1, ..., m\}$ *such that if* $x_1(t_0) = a_i$*,* $x_2(t_0) = b_i$*,* $s_{1i} = e_{1k}$*, and* $s_{2i} = e_{2k'}$*, then* $x_1(t) = a_{i+1}$ *and* $x_2(t) = b_{i+1}$ *in formula (3.7).*

ii) For the vertex (a_L, b_L)*, there exist* $t > t_0$*,* $k \in \{1, ..., n\}$*,* $k' \in \{1, ..., m\}$ *such that if*

$x_1(t_0) = a_L$, $x_2(t_0) = b_L$, $s_{1L} = e_{1k}$, and $s_{2L} = e_{2k'}$, then $x_1(t) = a_1$ and $x_2(t) = b_1$ in formulae (3.7).

3.3 The Existence of a Shapley Polygon as a Sufficient Condition for the Existence of a Nash Equilibrium

This section delivers a proof that the existence of a Shapley polygon is a sufficient condition for the existence of a Nash equilibrium. The proof follows from a matrix reduction technique described in the following Theorem.

Theorem 3.1 *Let a bimatrix game be defined by the matrices $A_{n\times m}$ and $B_{m\times n}$. Assume that fictitious play initialized at $(x_1(0), x_2(0))$ converges to a Shapley polygon which can be identified with the set of vertices $V = \{(a_l, b_l)|a_l \in \Delta_{x_1}, b_l \in \Delta_{x_2}\}_{l=1}^{L}$ and attractors $S = \{(s_{1l}, s_{2l})|s_{1l} \in E_1, s_{2l} \in E_2\}_{l=1}^{L}$ (see Definition 3.2). Furthermore, assume that the sets S_a and S_b are defined as the sets of all of the attractors for the vertices of V, i.e.,*

$$S_a \triangleq \{i \in \{1, ..., n\}|\exists(s_{1l}, s_{2l}) \in S, s_{1l} = e_{1i}\} \tag{3.8}$$

$$S_b \triangleq \{i \in \{1, ..., m\}|\exists(s_{1l}, s_{2l}) \in S, s_{2l} = e_{2i}\}. \tag{3.9}$$

Let matrix A' be a submatrix of A resulted by eliminating every row $i \in \{1, ..., n\}$ in A if $i \notin S_a$ and eliminating every column $j \in \{1, ..., m\}$ in A if $j \notin S_b$. Also, let matrix B' be a submatrix of B resulted by eliminating every row $i \in \{1, ..., m\}$ in B if $i \notin S_b$ and eliminating every column $j \in \{1, ..., n\}$ in B if $j \notin S_a$. Furthermore, let I_A and I_B be vectors such that their i-th component is equal to their row number in the reduced matrices A' and B' respectively, and otherwise zero, e.g., if A has six rows and $Sa = \{1, 2, 6\}$, then $I_A = [1, 2, 0, 0, 0, 3]$.

Then, there exist mixed strategies $\tilde{x}_1 \in \Delta_{n'} \triangleq \{\zeta \in R_+^{n'}|\Sigma_{i=1}^{n'}\zeta_i = 1\}$ and $\tilde{x}_2 \in \Delta_{m'} \triangleq \{\zeta \in R_+^{m'}|\Sigma_{i=1}^{m'}\zeta_i = 1\}$, and vectors $\Gamma_A = \gamma_A\mathbf{1}$, and $\Gamma_B = \gamma_B\mathbf{1}$, where $\mathbf{1}$ is the vector of appopriate size with all components equal to 1, γ_A and $\gamma_B \in R$, n' is the cardinality of S_a and m' is the cardinality of S_b, then matrices $A_{m\times n}$ and $B_{n\times m}$ reduce to matrices $A'_{n'\times m'}$ and

$B'_{m'\times n'}$, *such that* $A'\tilde{x}_2 = \Gamma_A$ *and* $B'\tilde{x}_1 = \Gamma_B$. *Henceforth,* $(\tilde{x}_1, \tilde{x}_2)$ *is a Nash equilibrium for the bimatrix game defined by the matrices* A' *and* B'.

Moreover, if the vectors $\hat{x}_1 \in R^n$ *and* $\hat{x}_2 \in R^m$ *are defined by*

$$\hat{x}_1^i = \begin{cases} \tilde{x}_1^{I_A^i} & i \in S_a \\ 0 & i \notin S_a \end{cases} \tag{3.10}$$

$$\hat{x}_2^i = \begin{cases} \tilde{x}_2^{I_B^i} & i \in S_b \\ 0 & i \notin S_b \end{cases} \tag{3.11}$$

then $(\hat{x}_1, \hat{x}_2)$ *is a Nash equilibrium of the bimatrix game defined by matrices* A *and* B.

Proof. It follows from (3.8) that for every $i \in S_a$ there exist a pair of attractors $(s_{1l}, s_{2l}) \in S$ such that $s_{1l} = e_{1i}$ (see (3.4)). In other words, there exist $\beta_i \in \Delta_{x_2}$ such that $(A\beta_i)^i \geq (A\beta_i)^{i'}$ for every $i' \in S_a$. For the simplicity of notation, let $\zeta_i \triangleq (A\beta_i)^i$. Since the set S_a is finite, there exist $i_l, i_u \in S_a$ such that for every $i \in S_a$, $\zeta_{i_l} \leq \zeta_i \leq \zeta_{i_u}$. Let $\beta_{i_l}, \beta_{i_u} \in \Delta_{x_2}$ be chosen such that $\zeta_{i_l} = (A\beta_{i_l})^{i_l}$ and $\zeta_{i_u} = (A\beta_{i_u})^{i_u}$.

Let matrix A'' be defined by row relocation in matrix A' such that if $k > k'$ and $\zeta_k > \zeta_{k'}$, then the rows k and k' interchange. Define the mapping $f : \Delta_{x_2} \mapsto Y \subset R^{n'}$ by $f(\lambda) = A''\lambda$. Let the lexicographical order for vectors $v_1, v_2 \in R^{n'}$ be defined by: $v_1 \preceq v_2$ if and only if there exists $k \in \{1, ..., n'\}$ such that $v_1^k \leq v_2^k$ and $v_1^{k'} = v_2^{k'}$ for every $k' \in \{1, ..., n'\}$, if $k' < k$.

Since $\zeta_{i_l} \leq \zeta_{i_u}$, there exists $\gamma_A \in R$ such that $\zeta_{i_l} \leq \gamma_A \leq \zeta_{i_u}$. Consequently, there exists $\Gamma_A = \gamma_A \mathbf{1}$, such that $A''\beta_{i_l} \preceq \Gamma_A \preceq A''\beta_{i_u}$, where $\preceq$ signifies the lexicographical order introduced above. The Intermediate Value Theorem [113] is used to prove that there exists a vector $\tilde{x}_2 \in \Delta_{m'}$ such that $A''\tilde{x}_2 = \Gamma_A$. In other words, there exists a vector in the range of $f(\lambda) = A''\lambda$ such that all of its components are equal. To prove this claim, the following facts are considered:

1) $\Delta_{m'}$ is a connected set,

2) by the lexicographical order considered above, $A''\beta_{i_l} \preceq \Gamma_A \preceq A''\beta_{i_u}$,

3) $f(\zeta) = A''\zeta$ is a continuous mapping.

Also, one can conclude that $A'\tilde{x}_2 = \Gamma_A$, since A'' is obtained by row relocation in A' and the fact that all the components of Γ_A are equal. By a similar argument, there exists $\tilde{x}_1 \in \Delta_{n'}$, $B'' \in R^{m' \times n'}$ which is obtained by row relocation in B', and $\Gamma_B = \gamma_B \mathbf{1}$, such that $B''\tilde{x}_1 = \Gamma_B$. Also, $B'\tilde{x}_1 = \Gamma_B$, since B'' is obtained by row relocation in B', and since all the components of Γ_B are equal. The pair $(\tilde{x}_1, \tilde{x}_2)$ is a Nash equilibrium of the bimatrix game defined by matrices A' and B' because:

i) $A'\tilde{x}_2 = \Gamma_A$, thus for every $x'_1 \in \Delta_{n'}$, $u_1(x'_1, \tilde{x}_2) = \gamma_A$, and there exists no $x'_1 \in \Delta_{n'}$ such that $u_1(x'_1, \tilde{x}_2) > u_1(\tilde{x}_1, \tilde{x}_2)$

ii) $B'\tilde{x}_1 = \Gamma_B$, thus for every $x'_2 \in \Delta_{m'}$, $u_2(\tilde{x}_1, x'_2) = \gamma_B$, and there exists no $x'_2 \in \Delta_{m'}$ such that $u_2(\tilde{x}_1, x'_2) > u_2(\tilde{x}_1, \tilde{x}_2)$.

Finally, to prove that $(\hat{x}_1, \hat{x}_2)$ is a Nash equilibrium for the bimatrix game defined by matrices A and B, the components of vectors $B\hat{x}_1$ and $A\hat{x}_2$ are considered. Using (3.10), (3.11), elimination rule, and the equations $A'\tilde{x}_2 = \Gamma_A$, $B'\tilde{x}_1 = \Gamma_B$,

$$(B\hat{x}_1)^i = \begin{cases} \gamma_A & i \in S_a \\ \gamma'_{Ai} & i \notin S_a \end{cases} \tag{3.12}$$

$$(A\hat{x}_2)^i = \begin{cases} \gamma_B & i \in S_b \\ \gamma'_{Bi} & i \notin S_b \end{cases} \tag{3.13}$$

where $\gamma'_{Ai} < \gamma_A$ and $\gamma'_{Bi} < \gamma_B$, because the rows containing γ'_{Ai} and γ'_{Bi} have been eliminated in the process of matrix reduction. Consequently, there exists no $x_1 \in \Delta_{x_1}$ such that $x_1^T A\hat{x}_2 > \hat{x}_1^T A\hat{x}_2$, and there exist no $x_2 \in \Delta_{x_2}$ such that $x_2^T B\hat{x}_1 > \hat{x}_2^T B\hat{x}_1$. Therefore, $(\hat{x}_1, \hat{x}_2)$ is a Nash equilibrium for the bimatrix game defined by matrices A and B.

Corollary 3.1 *To facilitate the procedure of finding a solution to the equation* $A'\tilde{x}_2 = \Gamma_A$, *where* $\tilde{x}_2 \in \Delta_{m'}$ *and* Γ_A *is defined in the proof of Theorem 3.1, one can solve the equation*

$\Lambda_A \tilde{x}_2 = \Theta_A$, where $\Lambda_A = [\lambda_A^{ij}]_{n' \times m'}$ is defined by

$$\lambda_A^{ij} = \begin{cases} 1 & i = 1 \\ a'^{1j} - a'^{ij} & i = \{2, ..., n'\}, \end{cases} \tag{3.14}$$

$\Theta_A \in \{R^{n'} | \Theta_A^1 = 1, \Theta_A^i = 0, \forall i \in \{2, ..., n'\}\}$, and restrict the answer to $\tilde{x}_2 \in \Delta_{x_2}$.

Similarly, to find $\tilde{x}_1 \in \Delta_{n'}$, such that $B' \tilde{x}_1 = \Gamma_B$, one can solve the equation $\Lambda_B \tilde{x}_1 = \Theta_B$, where $\Lambda_B = [\lambda_B^{ij}]_{n' \times m'}$ is defined by

$$\lambda_B^{ij} = \begin{cases} 1 & i = 1 \\ b'^{1j} - b'^{ij} & i = \{2, ..., m'\}, \end{cases}$$

$\Theta_B \in \{R^{m'} | \Theta_B^1 = 1, \Theta_B^i = 0, \forall i \in \{2, ..., m'\}\}$, and restrict the answers to $\tilde{x}_1 \in \Delta_{x_1}$.

Proof. It follows from (3.14) that for every $\tilde{x}_2 \in \{R^{m'} | \sum_{i=1}^{m'} \tilde{x}_2^i = 1\}$, $A' \tilde{x}_2 = \Gamma_A$ if and only if $\Lambda_A \tilde{x}_2 = \Theta_A$. From Theorem 3.1, there exists $\tilde{x}_2 \in \Delta_{x_2}$ such that $A' \tilde{x}_2 = \Gamma_A$, thus the set of solutions for equation $\Lambda_A \tilde{x}_2 = \Theta_A$, where $\tilde{x}_2 \in \Delta_{x_2}$ is a nonempty set which can also be used as the set of solutions for the equation $A' \tilde{x}_2 = \Gamma_A$, where $\tilde{x}_2 \in \Delta_{x_2}$. The proof for the second part of Corollary 3.1 is similar.

Proposition 3.2 *For every row $i \in \{1, ..., n\}$ in matrix A, $i \notin S_a$, where S_a is defined in (3.8), if and only if there exists a vertex (a_l, b_l) of the Shapley polygon, where the i-th component of the vector a_l, $a_l^i = 0$.*

For every row $i \in \{1, ..., m\}$ in matrix B, $i \notin S_b$, where S_b is defined in (3.9), if and only if there exists a vertex (a_l, b_l) of the Shapley polygon, where the i-th component of the vector b_l, $b_l^i = 0$.

Proof. For the Shapley polygon described in Definition 3.3 by the set of vertices $V = \{(a_l, b_l) | a_l \in \Delta_{x_1}, b_l \in \Delta_{x_2}\}_{l=1}^{L}$ and their corresponding attractors $S = \{(s_{1l}, s_{2l}) | s_{1l} \in E_1, s_{2l} \in E_2\}_{l=1}^{L}$ and for the row $i \in \{1, ..., n\}$ in matrix A, the i-th row of (3.7), for every vertex $(a_l, b_l), l \in \{1, ..., L-1\}$ is written as

$$a_{l+1}^i = (1 - \alpha_l) s_{1l}^i + \alpha_l a_l^i \tag{3.15}$$

where the variable $\alpha_l = \frac{(t_l - t_{l-1})}{t_l}$ depends on the two instances of time t_l and t_{l+1} in one complete round of a Shapley polygon, in which $x_1(t_l) = a_{l+1}$, $x_1(t_{l-1}) = a_l$, and $x(0) = a_1$.

For the vertex (a_L, b_L) the equation is written as

$$a_1^i = (1 - \alpha_L)s_{1l}^i + \alpha_L a_L^i \tag{3.16}$$

where $\alpha_L = \frac{(t_L - t_{L-1})}{t_L}$, $x_1(t_{L-1}) = a_L$, and $x_1(t_L) = a_1$. The recursive solution of (3.15) and (3.16), for all $l \in \{1, ..., L\}$, is given as

$$a_l^i = \frac{1}{1 - \prod_{i=1}^{L} \alpha_i} \Big(\sum_{k=(l-2)^*,...,l} \Big(\prod_{k'=(k+1)^*,...,(l-1)^*} \alpha_{k'} \Big) s_{1k}^i + (1 - \alpha_{l-1}) s_{l-1}^i \Big) \tag{3.17}$$

where the notation k^* is used to show a recursive integer of length L and is defined as $k^* \triangleq k$ if $k \in \{1, ..., L\}$, $k^* \triangleq k - L$ if $k^> L$, and $k^* \triangleq k + L$ if $k < 1$.

Since for every $l \in \{1, ..., L\}$, $s_{1k}^i \geq 0$, and for every $k \in \{1, ..., L\}$, $\alpha_k > 0$, one can conclude that for an arbitrary $i \in \{1, ..., n\}$, $a_l^i = 0$ if and only if for every $l \in \{1, ..., L\}$, $s_{1l}^i = 0$. Equivalently, for the row $i \in \{1, ..., n\}$ in matrix A, $i \notin S_a$ if and only if there exists $l \in \{1, ..., L\}$ such that $a_l^i = 0$ (see (3.8)).

The proof for the second part of Proposition 3.2 is similar.

3.4 A Description for the Process of Calculation of a Nash Equilibrium as Suggested by Theorem 3.1

Given a particular Shapley polygon, the sets S_a and S_b hold the information which are needed to perform the matrix reduction. It follows from the definition of S_a, (3.8), that $i \in S_a$ if and only if there exist $(\bar{a}_l, \bar{b}_l) \in V$ and $(e_{1i}, s_{2l}) \in S$ such that the strategy e_{1i} is a best response of player 1 to the strategy $\bar{b}_l$ by player 2, or equivalently $(A\bar{b}_l)^i \geq (A\bar{b}_l)^{i'}$, for every $i' \in \{1, ..., n\}$. Consequently, if $(\hat{x}_1, \hat{x}_2)$ is a Nash equilibrium of the game defined by A and B and $\hat{x}_2$ is a linear combination of the vectors $\{b_l\}_{l=1}^{L}$, for every $i \notin S_a$, then $\hat{x}_1^i = 0$, because there exists $j' \in \{1, ..., n\}$ such that $(A\tilde{x}_2)^i < (A\tilde{x}_2)^{i'}$. Similarly, if $i \notin S_b$, and $\hat{x}_1$ is a linear combination of $\{a_l\}_{l=1}^{L}$, then $x_2^i = 0$. Considering the necessarily zero components in $\hat{x}_1$, $\hat{x}_2$ and the utility function in (3.2), it is explained why matrices A and B are reduced to A' and B' using S_a and S_b. Theorem 3.1 guarantees that a Nash equilibrium exists for the reduced game.

Proposition 3.2 helps with a quick determination of the sets S_a and S_b when the fictitious play has converged to a Shapley polygon, and Corollary 3.1 offers a practical way for finding the solution.

In the process of reducing matrix A to matrix A', if the i-th pure strategy of player 1, represented by the row i in matrix A is strictly dominated by its i'-th strategy, then the row i in matrix A will be eliminated. For every strategy of player 2, $\beta \in \Delta_{x_2}$, $(A\beta)^{i'} > (A\beta)^{i}$ holds, thus $i \notin S_a$. However, depending on which Shapley polygon is reached, some of the non-dominated strategies may also be eliminated in the reduction process. Similarly, the matrix B is reduced to B' by elimination of all of its strictly dominated strategies. As in the case for matrix A, the non-dominated strategies may also be eliminated.

3.5 Numerical Examples

Two numerical examples are presented here that show how a Nash equilibrium of the game is calculated by the proposed matrix reduction technique if fictitious play has converged to a Shapley polygon. The first example is generated by randomly choosing two 4×4 integer matrices, A and B, and two initial mixed strategies, $x_1(0)$ and $x_2(0)$. The second example is taken from [105], and was also presented as the motivational example in the beginning of this chapter.

Example 1: Let a bimatrix game be defined by the matrices

$$A = \begin{bmatrix} -4 & 0 & -2 & 1 \\ 2 & -3 & 2 & 3 \\ -5 & -3 & -2 & -4 \\ -1 & -1 & 4 & -5 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & -3 & -5 & -3 \\ -2 & 2 & 1 & -2 \\ 4 & 1 & 0 & -4 \\ 4 & -5 & 1 & -3 \end{bmatrix}$$

and let the initial strategies for learning by fictitious play be given as

$$x_1(0) = [0.341028, 0.05402, 0.134054, 0.470817]^T$$

$$x_2(0) = [0.27749, 0.31584, 0.23102, 0.17565]^T.$$

The game converges to a Shapley polygone which induces the supports

$$S_a = \{1, 2, 4\}, \quad S_b = \{2, 3, 4\},$$

and by the matrix reduction introduced in Theorem 3.1, the matrices A' and B' are obtained as

$$A' = \begin{bmatrix} 0 & -2 & 1 \\ -3 & 2 & 3 \\ -1 & 4 & -5 \end{bmatrix}, \qquad B' = \begin{bmatrix} -2 & 2 & -2 \\ 4 & 1 & -4 \\ 4 & -5 & -3 \end{bmatrix}.$$

Following Corollary 3.1, the equations to solve are $\Lambda_A \tilde{x}_2 = \Theta_A$ and $\Lambda_B \tilde{x}_1 = \Theta_B$ where

$$\Lambda_A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -4 & -2 \\ 1 & -6 & 6 \end{bmatrix}, \qquad \Lambda_B = \begin{bmatrix} 1 & 1 & 1 \\ -6 & 1 & 2 \\ -6 & 7 & 1 \end{bmatrix}$$

and

$$\Theta_A = \Theta_B = [1, 0, 0]^T.$$

Nash equilibrium, $(\tilde{x}_1, \tilde{x}_2)$, for the bimatrix game defined by A' and B' is calculated as $\tilde{x}_1 = [0.2364, 0.1091, 0.6545]^T$ and $\tilde{x}_2 = [0.5143, 0.2857, 0.2]^T$. Consequently, using (3.12) and (3.13), the equilibrium for the bimatrix game defined by matrices A and B, $(\hat{x}_1, \hat{x}_2)$, is calculated, and can be verified by calculating Γ_A and Γ_B, as $\hat{x}_1 = [0.2364, 0.1091, 0, 0.6545]^T$ and $\hat{x}_2 = [0, 0.5143, 0.2857, 0.2]^T$.

Example 2: The motivation example from [105] which was presented in the beginning of this Chapter is considered. This one population example is equivalent to a bimatrix game in which both players have the same utility matrix, (see 3.1), and the same initial mixed strategies.

As illustrated by [105], the continuous time fictitious play converges to a Shapley polygon. In [105], the Time Average of the Shapley Polygon (TASP) is introduced to calculate

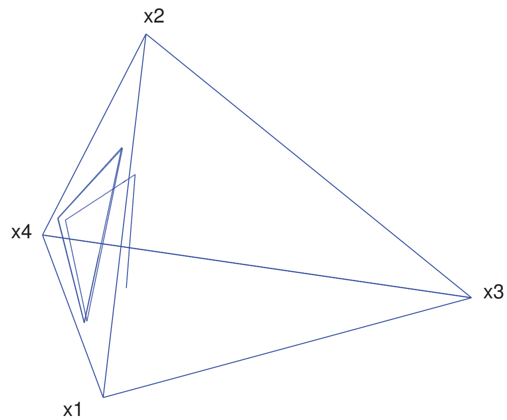


Figure 3–1: Mixed strategy polytope for player 1. The vertices X1, X2, X3 and X4 represent the pure strategies $[1, 0, 0, 0]^T$ to $[0, 0, 0, 1]^T$, respectively.

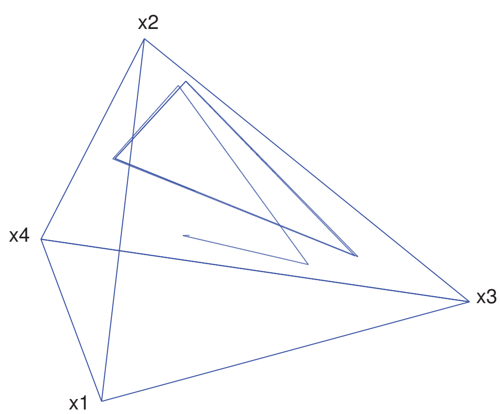


Figure 3–2: Mixed strategy polytope for player 2. The vertices X1, X2, X3 and X4 represent the pure strategies $[1, 0, 0, 0]^T$ to $[0, 0, 0, 1]^T$, respectively.

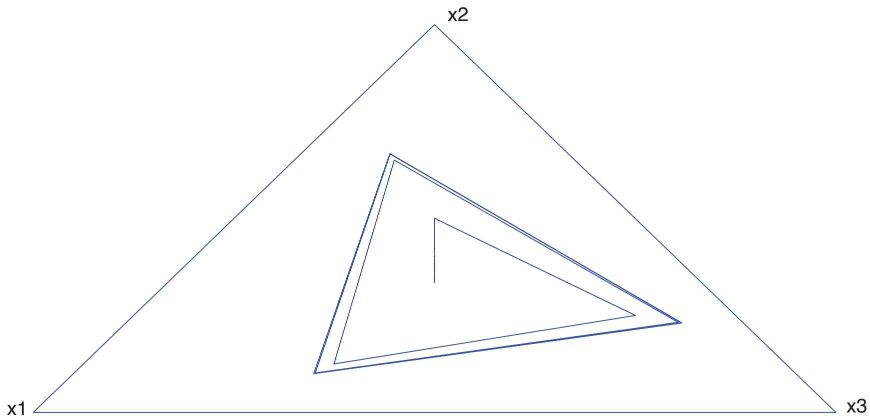


Figure 3–3: Convergence of fictitious play to a Shapley polygon for both players, for the initial mixed strategies $x(0) = [1/3, 1/3, 1/3]$. The vertices X1, X2 and X3 represent the pure strategies $[1, 0, 0]^T$, $[0, 1, 0]^T$ and $[0, 0, 1]^T$, respectively.

only an estimate for a Nash equilibrium. However, this estimate can be distinct from a Nash equilibrium for some games [105]. Using Proposition 3.2 and Theorem 3.1 the exact Nash equilibrium can be found. In this case, the matrix A can not be reduced and by Corollary 3.1, matrix Λ_A is calculated as

$$\Lambda_A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ -3 & 1 & 0 \end{bmatrix}.$$

By solving the equation $\Lambda_A \hat{x} = \Theta_A$, where $\Theta_A = [1, 0, 0]^T$ a Nash equilibrium of the game is calculated as $\hat{x} = [9/32, 10/32, 13/32]$.

3.6 Summary

Fictitious play was originally designed to calculate Nash equilibria in a bimatrix game [26]. However, Shapley showed that fictitious play may not converge to a Nash equilibrium but to a polygon [30]. In the present thesis, it is first proven that in a bimatrix game, existence of a Shapley polygon is the sufficient condition for the existence of a Nash equilibrium; this fact is then used to calculate a Nash equilibrium of the game. Contrary to the time averaging method proposed by [105], the Nash equilibrium calculated by the proposed method is exact.

CHAPTER 4
Stochastic Learning by Following the Best Bidder in Double Auction Markets

Auctions are the basic framework of exchange in many markets and can be analyzed by game theoretical methods. In any game, players attempt to maximize their individual utility function. If the game is repetitive, the players can use the past experience to learn to play better in the future. It is then of primary interest to determine whether the repeated game can converge to some kind of equilibrium. Learning and convergence are particularly difficult when utility functions exhibit discontinuities as is the case of utility functions used in auction markets.

In this chapter, an algorithm for learning in a double auction market is proposed in which a population of buyers and a population of sellers trade. Buyers and sellers know all the bids of the previous round but they do not know the bids of any other players in the current round. All the buyers bid as much as the most successful buyer in the previous round plus a Gaussian mutation that diminishes as the game progresses. Also, all the sellers bid as much as the most successful seller in the previous round plus a diminishing Gaussian mutation. Evolutionary algorithms which use the concept of simulated annealing [114], share the same principle of a diminishing probability in making a suboptimal choice.

The objective of writing this algorithm is not to calculate a Nash equilibrium of the game which is non-unique and can be determined by easier ways but to know if convergence to a Nash equilibrium can be guaranteed when the players update their strategies in a certain way.

4.1 Problem Statement

A double auction market is considered in which the number of buyers and sellers is the same and is equal to $n > 0$. It is assumed that in any round of the game every buyer has the possibility of trading with m randomly chosen sellers. A transaction will take place, benefiting both the buyer and the seller, only if the price bid by the buyer exceeds the price asked by the seller. The utility function for all buyers is defined as

$$u_b(x, y) : [c, v] \times [c, v] \to \Re \tag{4.1}$$

and the utility function for all sellers is defined as

$$u_s(x, y) : [c, v] \times [c, v] \to \Re \tag{4.2}$$

where c is the minimum price that a seller may trade, v is the value of the object for buyers, x and y are the prices bid and offered by the buyer i, and the seller j who is randomly matched with this buyer. The following assumptions are made for the utility function of buyers and sellers:

A1) $u_b(x, y) = 0$ and $u_s(x, y) = 0$, if $x < y$.

A2) The utility function for buyers, $u_b(x, y)$ and the utility function for sellers $u_s(x, y)$ are twice differentiable, i.e., for every $x \in [c, v]$ and $y \in [c, x]$, $\frac{\partial^2 u_b(x,y)}{\partial x^2}$ and $\frac{\partial^2 u_s(x,y)}{\partial y^2}$ are defined[1] .

A3) for all $x \in [c, v]$ and $y \in [c, x]$, $\frac{\partial u_b(x,y)}{\partial x} \leq 0$ and $\frac{\partial u_s(x,y)}{\partial y} \geq 0$.

Proposition 4.1 *(Similar to [91]) A double auction game employing any utility function that satisfies the assumptions A1, A2 and A3 with populations of buyers and sellers of equal cardinalities, is in equilibrium if all the players are bidding and offering the same price, i.e.* $p_{bi} = p_{sj} = \psi \in [c, v]$ *for all* $i, j \in \{1, ..., n\}$.

Proof. Assume that all the players are bidding $\psi \in [c, v]$ i.e. $p_{bi} = p_{sj} = \psi, \forall i, j$. If bidder i decides to bid higher, $p_{bi} > \psi$, because of the assumption $A3$, the utility function will decrease for that player. If the same bidder decides to bid lower, his/her utility function

[1] Definition of the differentiation is restricted to the interval $D \triangleq \{(x, y)|x \in [c, v], y \in [c, v], y \leq x\}$.

will be zero. A similar reasoning can be applied to a seller, implying that the market is in a Nash equilibrium.■

The problem is stated as whether this game, defined with two utility functions $u_b(.,.)$ and $u_s(.,.)$ which satisfy all the above assumptions, will converge to a Nash equilibrium of the game under the learning procedure stated in the Algorithm 4.1, explained below.

4.2 Algorithm Proposed for Learning in Double Auction Markets

The evolutionary iterative algorithm for learning in double auction markets developed here belongs to the general class of random search algorithms. The underlying idea of the algorithm is that buyers and sellers try to follow the most successful buyer or seller known to them from the previous iteration of the algorithm. The following notation is adopted :

– n, the cardinality of the populations of buyers and sellers.

– k, the index of the current round of the game, ($k \in \mathbb{Z}$.)

– $p_{bi}(k) \in [c, v]$, the maximal price at which buyer i is willing to buy in round k.

– $p_{sj}(k) \in [c, v]$, the minimal price at which seller j is willing to sell in round k.

– $c \in [0, 1]$, the cost of production for sellers.

– $v \in [0, 1]$, the value of the product for buyers.

– $\bar{p}_b(k) \in [c, v]$, the average of the buyers' bid prices in round k, i.e., $\bar{p}_b(k) = \sum_{i=1}^{n} p_{bi}(k)/n$.

– $\bar{p}_s(k) \in [c, v]$, the average of the sellers' ask prices in round k, i.e., $\bar{p}_s(k) = \sum_{j=1}^{n} p_{sj}(k)/n$.

– m, the number of sellers that any buyer will be matched to in any round of the game.

– $\alpha > 1$, a shrinking factor for the variance of the randomizer function used in the generation of the bid and offer prices.

– $N(\mu, \sigma)$, the normal distribution with mean μ and variance σ.

– $\sigma_k > 0$, the variance of the random generator function in round k of the game.

– $\mu_{bk} \in [c, v]$ and $\mu_{sk} \in [c, v]$, the mean values for the normal probability distribution functions for buyers and sellers, respectively.

– m_{count}, a counter by which a buyer will be matched to exactly m sellers.

– $usum_{bi}$, variable that is used in averaging the utilities of buyers. (See Step 4)

– $usum_{sj}$, variable that is used in averaging the utilities of sellers. (See Step 4)

– cs_j, the counter of the number of times that seller j has a chance to participate in a transaction.

– $\epsilon \in (0, 1)$, algorithm termination threshold.

– i^* and j^*, the indices of the buyer and seller, respectively, who achieve the highest utility values in the current round of the game.

– $\hat{j}$, the index of the seller who is randomly chosen to match the buyer i.

Before stating the steps of the algorithm, it is helpful to explain the meaning behind them. The values of the algorithm parameters and the initial values of the buyers' and sellers' prices are selected in Steps 0 and 1. The latter are variables that are used in averaging the utility of every buyer and seller that participate in the market. Steps 3 - 6 constitute a loop in which each buyer is matched with m sellers in the current round of the algorithm. As a result of the matching between buyer i and seller j, both of them claim utility values $u_{bi}(p_{bi}, p_{sj})$, and $u_{sj}(p_{bi}, p_{sj})$, that add up to: $usum_{bi}$ and $usum_{sj}$, respectively. The counter cs_j is incremented to serve the averaging of utility values for every seller in Step 7. Buyers do not need a similar counter as there is always m values to average over for each buyer. Step 8, commences by determining the indices i^* and j^* of the buyer and seller, respectively, who achieve the highest utility values in the current round of the game. The prices of this buyer and seller are then selected as the averages μ_{bk} and μ_{sk} for the randomizer normal distribution employed to generate the prices for buyers and sellers in the next round of the game.

The variances of both probability distributions are shrunk by a factor $1/\alpha$ for the next round of the game. The variances of the randomizing distributions decrease as players learn about the market whose behaviour is tightly related to the ensemble of players' utility functions. The algorithm is exited if the prices of the buyers and sellers are sufficiently close to each other (close to the equilibrium of the game). Clearly, the information structure in this game is as follows: the players know their own utility functions, their own price, and the current price of their opponents in the market game.

Algorithm 4.1 Learning by Following the Best in Double Auction Markets.

Step 0: Set the initial values of m, c, v, $\alpha > 1$, *and* $k = 0$. *Set initial values for* $\sigma_0 > 0$, $\mu_{b0} \in [c, v]$ *and* $\mu_{s0} \in [c, v]$. *For* $i, j = \{0, ..., n\}$ *draw samples of initial values of the ask and bid prices from uniform distributions over the interval* $[c, v]$, *i.e.* $p_{bi}(0) \sim U(c, v)$ *and* $p_{sj}(0) \sim U(c, v)$.

Step 1: For $i, j = \{1, ..., n\}$ *set* $usum_{bi} = 0$, $usum_{sj=0} = 0$, *and* $cs_j = 0$ *(parameters needed for averaging the utility values for all players).*

Step 2: Set $i = 1$, *indicating that the utility function is averaged for buyer* i. *Set* $m_{\text{count}} = 0$.

Step 3: Draw an integer $\hat{j} \in \{1, 2, ..., n\}$ *from a uniform distribution (i.e.* $Pr(\hat{j}) = 1/n$*).*

Step 4: Calculate $u_b(i, j)$ *and* $u_s(i, j)$ *- the utility values for buyer* i *and seller* j *as they match. Update the total sums:* $usum_{bi} = usum_{bi} + u_{bi}(p_{bi}, p_{sj})$ *and* $usum_{sj} = usum_{sj} + u_{sj}(p_{bi}, p_{sj})$. *Update the counter of the number of times that seller* $\hat{j}$ *participates in asking against all buyers:* $cs_{\hat{j}} = cs_{\hat{j}} + 1$.

Step 5: Increment counter $m_{\text{count}} = m_{\text{count}} + 1$. *If* $m_{\text{count}} < m$, *go to Step 3.*

Step 6: $i = i + 1$, *go to Step 3 if* $i < n + 1$.

Step 7: For $i, j = \{1, ..., n\}$, *set* $u_{bi} = usum_{bi}/m$, *and* $u_{sj} = usum_{sj}/cs_j$, *the average utilities of buyers and sellers.*

Step 8: Update the price generator densities for the buyers and the sellers, as follows. First determine the indices i^* *and* j^* *of the buyer and seller, respectively, who achieve the highest utility values in the current round of the game. Then set :* $\mu_{bk} = p_{bi^*}$, *and* $\mu_{sk} = p_{sj^*}$.

Step 9: Evolve the price of each buyer and seller according to $p_{bi}(k + 1) \sim N(\mu_{bk}, \sigma_k)$, $p_{sj}(k + 1) \sim N(\mu_{sk}, \sigma_k)$; $i, j \in \{1, ..., n\}$. *If* $p_{bi} < c$, *then* $p_{bi} = c$. *If* $p_{bi} > v$, *then* $p_{bi} = v$. *If* $p_{sj} < c$, *then* $p_{sj} = c$. *If* $p_{sj} > v$, *then* $p_{sj} = v$.

Step 10: Update the variance σ_k, *e.g.,* $\sigma_{k+1} = \sigma_k/\alpha$.

Step 11: Verify the algorithm's stopping condition. If $|\bar{p}_b(k) - \bar{p}_s(k)| > \epsilon$, or $\sigma_k > \epsilon$ then set $k = k + 1$, and go to Step 1, or else exit the algorithm.

4.3 Proof of Convergence

Theorem 4.1 *Assume learning in a double auction by the above mentioned algorithm together with these conditions:*

1) utilities of buyers and sellers satisfy assumptions A1, A2, and A3.

2) buyers and sellers are risk averse or risk neutral, i.e.,

$$\frac{\partial^2 u_b(p_b, p_s)}{\partial p_b^2} \leq 0 \tag{4.3}$$

$$\frac{\partial^2 u_s(p_b, p_s)}{\partial p_b^2} \leq 0 \tag{4.4}$$

where $u_b(p_b, p_s)$ and $u_s(p_b, p_s)$ are the utility functions for all the buyers and all the sellers, respectively.

3) the first derivatives of utility functions $u_b(p_b, p_s)$ and $u_s(p_b, p_s)$ are bounded [2], i.e.,

$$\lambda_b \leq \frac{\partial u_b(p_b, p_s)}{\partial p_b} \leq 0 \quad \textit{and} \quad \lambda_s \leq \frac{\partial u_b(p_b, p_s)}{\partial p_s} \leq 0 \tag{4.5}$$

$$0 \leq \frac{\partial u_s(p_b, p_s)}{\partial p_b} \leq \lambda_b' \quad \textit{and} \quad 0 \leq \frac{\partial u_s(p_b, p_s)}{\partial p_s} \leq \lambda_s' \tag{4.6}$$

4) random generators used in the algorithm do not use a variance of zero.

In case these four assumptions are satisfied, then for every $\epsilon > 0$ there exist n_0, $m_0 \in \mathbb{N}$ and a sequence of variances, $\{\sigma_k > 0\}_{k=1}^{\infty}$, such that for every $n > n_0$ (number of buyers and sellers) and $m > m_0$ (number of times they meet); the proposed algorithm for learning in double auction markets, Algorithm 4.1, converges in probability to a Nash equilibrium. That is, $\lim_{k\to\infty} Pr(|\mu_{bk} - \mu_{sk}| > \epsilon) = 0$, where μ_{bk} and μ_{sk} are the mean values of the Gaussian random generator functions $N(\mu_{bk}, \sigma_k)$ and $N(\mu_{sk}, \sigma_k)$ for buyers and sellers, at time k.

[2] Assumptions which are possible if trading happens at a midpoint for the prices of the buyer and the seller, e.g., (4.51) and (4.52).

Proof : In round k of the game, $\hat{\mu}_{b(k+1)}$ and $\hat{\mu}_{s(k+1)}$ are defined as

$$\hat{\mu}_{b(k+1)} \triangleq \arg\max_{\lambda \in [c,v]} \int_{c}^{\lambda} f_s(p_s) u_b(\lambda, p_s) dp_s \tag{4.7}$$

and

$$\hat{\mu}_{s(k+1)} \triangleq \arg\max_{\lambda \in [c,v]} \int_{\lambda}^{v} f_b(p_b) u_s(p_b, \lambda) dp_b \tag{4.8}$$

where $f_s(.)$, and $f_b(.)$, are the two Gaussian distributions from the algorithm.

The proof starts with showing that for every step of the game, k, and for every$\sigma_k \in]0, \Sigma_b[$ where

$$\Sigma_b = -\frac{\sqrt{2} u_b(v, v)}{\sqrt{\pi} \lambda_b} \tag{4.9}$$

$LB(\sigma_k)$ a lower bound for $\hat{\mu}_{b(k+1)}$ is calculated as the following

$$\hat{\mu}_{b(k+1)} \geq LB(\sigma_k) \triangleq \hat{\mu}_{sk} + \sqrt{2\sigma_k^2 \ln(\frac{\Sigma_b}{\sigma_k})}. \tag{4.10}$$

To this end, let

$$g_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) = \int_{c}^{x} f_s(p_s) u_b(x, p_s) dp_s. \tag{4.11}$$

The first derivative of $g_{u_b}(x)$ is calculated using the Leibniz's integral rule as

$$\frac{\partial}{\partial x} g_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) = \frac{\partial}{\partial x} \int_{c}^{x} f_s(p_s) u_b(x, p_s) dp \tag{4.12}$$

$$= f_s(x) u_b(x, x) + \int_{c}^{x} f_s(p_s) \frac{\partial u_b(x, p_s)}{\partial x} dp_s.$$

The second derivative is calculated as

$$g''_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) = \frac{\partial^2}{\partial x^2} \int_{c}^{x} f_s(p_s) u_b(x, p_s) dp_s \tag{4.13}$$

$$= \frac{\partial f_s(x)}{\partial x} u_b(x, x) + f_s(x) \frac{\partial u_b(x, x)}{\partial x}$$

$$+ f_s(x) \frac{\partial u_b(x, p_s)}{\partial x}|_{p_s = x} + \int_{c}^{x} f_s(p_s) \frac{\partial^2 u_b(x, p_s)}{\partial x^2} dp_s.$$

Next, it is claimed that if the conditions in the theorem are satisfied, then for every $\sigma_k < \Sigma_b$; where Σ_b is defined in (4.9), the *argmax* set below is a singleton

$$\{\zeta^*\} = \arg\max_{x\in[c,v]} g_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) \tag{4.14}$$

where $\zeta^* \in [\hat{\mu}_{sk}, v]$. The proof follows by considering the signs of four terms of (4.13) and

$$f_s'(x) = \frac{-(x-\hat{\mu}_{sk})}{\sqrt{2\pi}\sigma_k^3} e^{\frac{-(x-\hat{\mu}_{sk})^2}{2\sigma_k^2}}$$

which leaves only two possible cases for the sign of $g''_{u_b}(x, \sigma_k, \hat{\mu_{sk}})$ in the interval $[c, \hat{\mu}_{sk}[$.

case1: If $g''_{u_b}(c, \sigma_k, \hat{\mu}_{sk}) \leq 0$ then for all $x \in [c, \hat{\mu}_{sk}[$ $g''_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) \leq 0$.

case2: If $g''_{u_b}(c, \sigma_k, \hat{\mu}_{sk})$ changes sign only once from positive to negative in the interval of $x \in [c, \hat{\mu}_{sk}[$, i.e., there exists $x^* \in [c, \hat{\mu}_{sk}[$ such that for every $x \in [c, x^*[$, $g''_{u_b}(p_{b(k+1)}, \sigma_k, \hat{\mu}_{sk}) > 0$ and for every $x \in]x^*, \hat{\mu}_{sk}[$ $g''_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) < 0$.

In both above cases, there exist no $x \in [c, \hat{\mu}_{sk}[$ such that $g'_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) = 0$, if the two following conditions are satisfied:

c1) $g'_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) > 0$ for $x = c$.

c2) $g'_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) > 0$ for $x = \hat{\mu}_{sk}$.

Satisfaction of (c1) is trivial and (c2) is satisfied if the following holds:

$$f_s(\hat{\mu}_{sk}) u_b(\hat{\mu}_{sk}, \hat{\mu}_{sk}) > -\int_c^{\hat{\mu}_{sk}} f_s(p_s) \frac{\partial u_b(x, p_s)}{\partial x} dp_s. \tag{4.15}$$

Due to (4.5), (4.15) can be reduced to

$$\frac{1}{\sqrt{2\pi}\sigma_k} u_b(\hat{\mu}_{sk}, \hat{\mu}_{sk}) > -\frac{\lambda_b}{2} \tag{4.16}$$

and invoking (4.5), (4.16) is satisfied for every $\sigma_k \in]0, \Sigma_b[$ where

$$\Sigma_b = -\frac{\sqrt{2} u_b(v, v)}{\sqrt{\pi}\lambda_b}. \tag{4.17}$$

Therefore, for every $\sigma_k \in]0, \Sigma_b[$ and for every $x \in [c, \hat{\mu}_{sk}]$

$$g'_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) > 0 \tag{4.18}$$

while for $x \in]\hat{\mu}_{sk}, v]$

$$g''_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) < 0. \tag{4.19}$$

This makes the proof of (4.14) complete, i.e., for every $\sigma_k < \Sigma_b$, where Σ_b is defined in (4.9), there exists no $\zeta^* \in [c, \hat{\mu}_{sk}[$ while there exists a unique $\zeta^* \in [\hat{\mu}_{sk}, v]$ such that (4.14) is satisfied.

Then, to establish a lower bound for $\hat{\mu}_{b(k+1)}$, it is claimed that for any $\sigma_k \in]0, \Sigma_b[$, $\hat{\mu}_{b(k+1)} < \nu_{bk+1}$ if

$$\hat{\mu}_{b(k+1)} = \arg\max_{x \in [c,v]} g_{u_b}(x, \hat{\mu}_{sk}, \sigma_k) = \arg\max_{x \in [c,v]} \int_c^x f_s(p_s) u_b(p_b, p_s) dp_s \tag{4.20}$$

where $u_b(p_b, p_s) = -\lambda_b p_b - \lambda_s p_s + K$, and

$$\nu_{b(k+1)} = \arg\max_{x \in [c,v]} g_{w_b}(x, \hat{\mu}_{sk}, \sigma_k) = \arg\max_{x \in [c,v]} \int_c^x f_s(p_s) w_b(x, p_s) dp_s \tag{4.21}$$

where $w_b(p_b, p_s)$ satisfies (4.5), K is a constant, and $u_b(c, c) = w_b(c, c)$. This claim is proven by first invoking (4.5)

$$\forall p_s \in [c, v] \ \frac{\partial u_b(x, p_s)}{\partial x} \leq \frac{\partial w_b(x, p_s)}{\partial x} \tag{4.22}$$

since $u_b(c, c) = w_b(c, c)$ one can write that

$$\forall x, p_s \in [c, v] \ u_b(x, p_s) \leq w_b(x, p_s) \tag{4.23}$$

so for every $x \in [\hat{\mu}_{sk}, v]$

$$\frac{\partial}{\partial x} g_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) \leq \frac{\partial}{\partial x} g_{w_b}(x, \sigma_k, \hat{\mu}_{sk}) \tag{4.24}$$

and hence $\hat{\mu}_{b(k+1)}, \nu_{b(k+1)} \in [\hat{\mu}_{sk}, v]$ and for every $x \in [\hat{\mu}_{sk}, v]$

$$g''_{u_b}(x, \sigma_k, \hat{\mu}_{sk}) < 0 \tag{4.25}$$

and

$$g''_{w_b}(x, \sigma_k, \hat{\mu}_{sk}) < 0. \tag{4.26}$$

Thus, it is concluded that $\hat{\mu}_{b(k+1)} \leq \nu_{b(k+1)}$.

Consequently, a lower bound for $\hat{\mu}_{b(k+1)}$ with the utility function given as $u_b(p_b, p_s) = -\lambda_b p_b - \lambda_s p_s + K$ may serve as a lower bound for any other utility function for the buyers. The lower bound is the zero of (4.12), hence it can be obtained by solving

$$\frac{1}{\sqrt{2\pi}\sigma_k}(-(\lambda_s + \lambda_b)x + K)e^{\frac{-(x-\hat{\mu}_{sk})^2}{2\sigma_k^2}} = \frac{\lambda_b}{2} \tag{4.27}$$

because $\hat{\mu}_{b(k+1)} \in [\hat{\mu}_{sk}, v]$ and for every $x \in [\hat{\mu}_{sk}, v]$ $g''_{u_b}(x, \sigma_k, \hat{\mu}_{s(k+1)}) < 0$. Furthermore, for every $p_b, p_s \in [c, v]$ $u_b(p_b, p_s) \geq u_b(v, v)$. Thus by replacing $(-(\lambda_s + \lambda_b)x + K)$ by $u_b(v, v)$ to find the lower bound :

$$LB(\sigma_k) = \hat{\mu}_{sk} + \sqrt{2\sigma_k^2 \ln(-\frac{\sqrt{2}u_b(v,v)}{\sqrt{\pi}\lambda_b\sigma_k})}. \tag{4.28}$$

The same arguments are used for sellers to show that if

$$0 \leq \frac{\partial u_s(p_b, p_s)}{\partial p_b} \leq \lambda'_b \quad \text{and} \quad 0 \leq \frac{\partial u_s(p_b, p_s)}{\partial p_s} \leq \lambda'_s \tag{4.29}$$

and

$$\Sigma_s = \frac{\sqrt{2}u_s(c,c)}{\sqrt{\pi}\lambda'_s} \tag{4.30}$$

then for every $\sigma_k \in]0, \Sigma_s[$ the following upper bound is guaranteed for $\hat{\mu}_{s(k+1)}$

$$UB(\sigma_k) = \hat{\mu}_{bk} - \sqrt{2\sigma_k^2 \ln(\frac{\sqrt{2}u_s(c,c)}{\sqrt{\pi}\lambda'_s\sigma_k})}. \tag{4.31}$$

If in step k, $\hat{\mu}_{bk} > \hat{\mu}_{sk}$, then by choosing $\sigma_k = \sigma^*$, $\sigma^* < min(\Sigma_b, \Sigma_s)$ the equations (4.28) and (4.31) together show that a contraction factor γ exists such that

$$|\hat{\mu}_{b(k+1)} - \hat{\mu}_{s(k+1)}| \leq \gamma |\hat{\mu}_{bk} - \hat{\mu}_{sk}| \tag{4.32}$$

where

$$\gamma \leq max(1 - \beta, 0) < 1 \tag{4.33}$$

and

$$\beta = \frac{\sqrt{2}\sigma^*}{v-c}\left(\sqrt{\ln(-\frac{\sqrt{2}u_b(v,v)}{\sqrt{\pi}\lambda_b\sigma^*})} + \sqrt{\ln(\frac{\sqrt{2}u_s(c,c)}{\sqrt{\pi}\lambda'_s\sigma^*})}\right). \tag{4.34}$$

To prove the possibility of contraction for a case where in step k, $\hat{\mu}_{bk} < \hat{\mu}_{sk}$, it is shown that

$$\lim_{\sigma_k \to 0^+} \hat{\mu}_{b(k+1)} = \hat{\mu}_{sk} \tag{4.35}$$

and

$$\lim_{\sigma_k \to 0^+} \hat{\mu}_{s(k+1)} = \hat{\mu}_{bk}. \tag{4.36}$$

Proof of (4.35) follows from the calculation of the limit of $\hat{\mu}_{b(k+1)}$

$$\lim_{\sigma_k \to 0^+} \hat{\mu}_{b(k+1)} = \lim_{\sigma_k \to 0^+} \operatorname*{arg\,max}_{b \in [c,v]} \int_c^b f_s(p_s) u_b(b,p_s) dp_s \tag{4.37}$$

$$= \lim_{\sigma_k \to 0^+} \operatorname*{arg\,max}_{b \in [c,v]} \int_c^b \delta(p_s - \hat{\mu}_{sk}) u_b(b,p_s) dp_s \tag{4.38}$$

where $\delta(.)$ is the delta dirac function (see page 132 in [115]), hence

$$\lim_{\sigma_k \to 0^+} \hat{\mu}_{b(k+1)} = \hat{\mu}_{sk}. \tag{4.39}$$

Similarly, one can prove (4.36).

If in step k, $\hat{\mu}_{bk} < \hat{\mu}_{sk}$ then because of (4.35) and (4.36) there exists a small $\sigma_k \in]0,\infty[$ that reverses the order, i.e., $\hat{\mu}_{b(k+1)} > \hat{\mu}_{s(k+1)}$, henceforth there exist $\gamma < 1$ (see (4.32)) such that

$$|\hat{\mu}_{b(k+2)} - \hat{\mu}_{s(k+2)}| \leq \gamma |\hat{\mu}_{bk} - \hat{\mu}_{sk}|. \tag{4.40}$$

Next it is shown that for every $\epsilon_1, \epsilon_2 > 0$, there exist $n_0, m_0 \in \mathbb{N}$, and $\sigma_0 > 0$ such that if $n > n_0$, $m > m_0$, and $\sigma_k < \sigma_0$ then

$$Pr(|\hat{\mu}_{b(k+1)} - \mu_{b(k+1)}| > \epsilon_1) < \epsilon_2 \tag{4.41}$$

$$Pr(|\hat{\mu}_{s(k+1)} - \mu_{s(k+1)}| > \epsilon_1) < \epsilon_2. \tag{4.42}$$

To prove 4.41, for the buyer i who bids p_{bi}, the probability

$$Pr(|u_{bi} - g_{ub}(p_{bi}, \sigma_k, \hat{\mu}_{sk})| > \epsilon_1) \tag{4.43}$$

is considered as a measure which shows the difference between u_{bi} which is stochastically obtained in the algorithm and $g_{ub}(.)$ which is a deterministic function. From Steps 7 and 4 of the algorithm, (4.43) is writen as

$$Pr(|\frac{\sum_{j_1}^{j_m} u_{b(p_{bi},p_{sj})}}{m} - g_{ub}(p_{bi}, \sigma_k, \hat{\mu}_{sk})| > \epsilon_1), \tag{4.44}$$

where j_1 to j_m are the indices for the sellers that this buyer was matched to (Step 2). using the definition (4.11), there exist $m_0 \in \mathbb{N}$ and $\sigma_0 > 0$ such that

$$\lim_{m\to\infty} Pr(|\frac{\sum_{j_1}^{j_m} u_{b(p_{bi},p_{sj})}}{m} - g_{ub}(p_{bi}, \sigma_k, \hat{\mu}_{sk})| > \epsilon_1) < \epsilon_2 \tag{4.45}$$

because the probability of matching is uniform and the probaility of drawing a sample from a Normal distribution is arbitrary low by choosing $\sigma_k > 0$, since $\mu_{bk} \in [c, v]$. Thus, truncation does not matter if (4.41) is satisfied.

On the other hand,

$$\lim_{n\to\infty} Pr(\forall i \in \{1, ..., n\}|p_{bi} - \hat{\mu}_{b(k+1)})| > \epsilon_1) = 0. \tag{4.46}$$

Since, bid prices for sellers are generated by a Gaussian with non-zero variance. Finally, in the proof of the theorem the argmax function for $g_{ub}(p_{bi}, \sigma_k, \hat{\mu})$ is a singleton with the value of $\mu_{b(k+1)}$. Thus (4.41) is proven. The proof for (4.42) is similar.

Now, for every $\epsilon_1, \epsilon_2 > 0$, there exist $n_0, m_0 \in \mathbb{N}$ and $\sigma_0 > 0$ such that (4.41) and (4.42) are satisfied. Besides, by using (4.34)

$$\beta_0 = \frac{\sqrt{2}\sigma_0}{v - c}(\sqrt{\ln(-\frac{\sqrt{2}u_b(v, v)}{\sqrt{\pi}\lambda_b\sigma_0})} + \sqrt{\ln(\frac{\sqrt{2}u_s(c, c)}{\sqrt{\pi}\lambda_s'\sigma_0})}). \tag{4.47}$$

and by (4.40)

$$|\hat{\mu}_{b(k+2)} - \hat{\mu}_{s(k+2)}| \le (1 - \beta_0)|\hat{\mu}_{bk} - \hat{\mu}_{sk}| \tag{4.48}$$

where $0 < \beta_0 < 1$. Thus, (using (4.41), (4.42), and (4.48)), if $|\hat{\mu}_{bk} - \hat{\mu}_{sk}| = d_k$, then

$$Pr(|\hat{\mu}_{b(k+2)} - \hat{\mu}_{s(k+2)}| > (1-\beta_0)d_k + 2\epsilon_1) < \epsilon_2 \tag{4.49}$$

and in $2l$, an even number, of steps

$$Pr(|\hat{\mu}_{b(k+2l)} - \hat{\mu}_{s(k+2l)}| > \ ((1-\beta_0)^{2l}d_k + 2l\epsilon_1)\ ) < \epsilon_2. \tag{4.50}$$

Thus, by defining $\epsilon_3 = (1-\beta_0)^{2l}d_k + 2l\epsilon_1)$, there exist $l_0 \in N$ such that for every $l > l_0$, (4.50) is satisfied. ■

4.4 Numerical Experiments

If $c \in [0,1]$ is the cost of the production and $v \in [0,1]$ represents the value of good for the buyers, and under the assumption that a buyer and a seller will benefit from their transaction equivalently, the utility functions, u_{bi} and u_{sj}, of buyer i and the seller j in a single round of the game can be given by the formulae below [92] [91]:

$$u_{bi}(b_i, s_j) = \begin{cases} v - \frac{p_{bi}+p_{sj}}{2} & \text{if } p_{bi} \in [p_{sj}, v] \\ 0 & \text{otherwise} \end{cases} \tag{4.51}$$

$$u_{sj}(b_i, s_j) = \begin{cases} \frac{p_{bi}+p_{sj}}{2} - c & \text{if } p_{sj} \in [c, p_{bi}] \\ 0 & \text{otherwise} \end{cases} \tag{4.52}$$

in which p_{bi} and p_{sj} denote the prices of buyer i and seller j, respectively.

Another example of a double auction market will be also considered that is created by adopting a different set of utility functions:

$$u_{bi}(b_i, s_j) = \begin{cases} \sqrt{v - \frac{p_{bi}+p_{sj}}{2}} & \text{if } p_{bi} \in [p_{sj}, v] \\ 0 & \text{otherwise} \end{cases} \tag{4.53}$$

$$u_{sj}(b_i, s_j) = \begin{cases} \sqrt{\frac{p_{bi}+p_{sj}}{2} - c} & \text{if } p_{sj} \in [c, p_{bi}] \\ 0 & \text{otherwise} \end{cases} \tag{4.54}$$

The evolution of the ask and bid prices of the players in the market during the first few rounds of the game are shown in Figure 4–1. It can be seen that the prices of both buyers and sellers are concentrating in the neighborhoods of their corresponding best bid or ask prices. In all these tests the parameters of the algorithm are set to $n = 100$, $m = 20$, $\alpha = 1.1$, $\sigma_0 = 0.3$, $c = 0$, $v = 1$.

The curves in Figure 4–2 represent the evolution of the average prices of the population of buyers and sellers during the game using the utility functions (4.51) to (4.54), respectively. It is seen that convergence to a Nash equilibrium of the game is achieved in each case. Also, in the Figure 4–2, statistics of convergence are shown for the algorithm that terminates after 100 rounds of the auction game. It is seen that the spread between the average bid and ask prices $p_b - p_s$ is marginally small which essentially demonstrates convergence to a single market price.

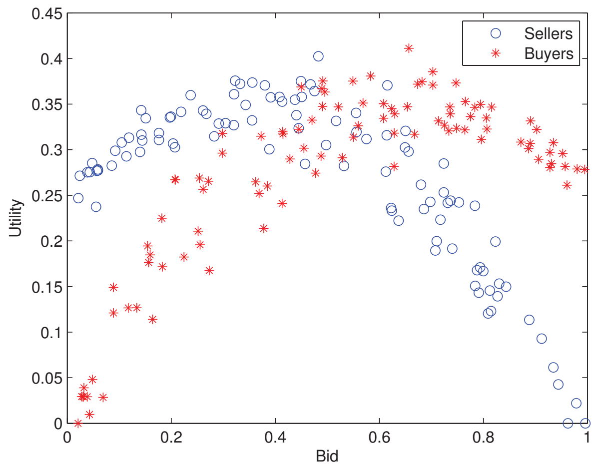
4.5 Merit Based Matching

In real auction markets the buyers and sellers do not match randomly, but the system selects the partners by their merits.

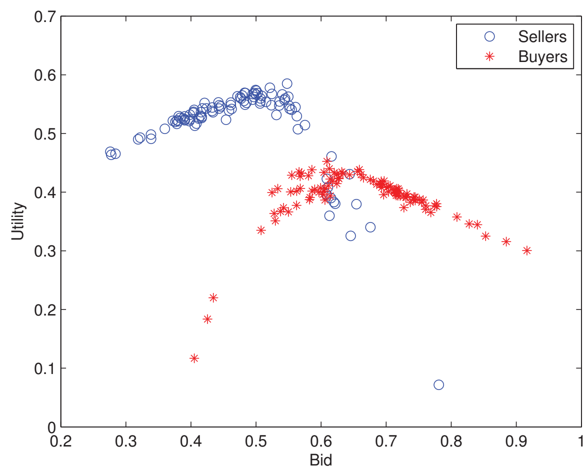
The main idea is that a buyer who bids a higher price has more merit for transaction than a buyer who bids a lower price and a seller who asks a lower price has more merit than a seller who asks a higher price. The buyer and the seller with highest merits are first matched and then other buyers and sellers are matched based on their ranks of merit. Algorithm 4.2 is designed for learning in this type of market. Unless otherwise stated, notations and definitions are identical to those of Algorithm 4.1.

Algorithm 4.2 Algorithm for Learning in Double Auction Markets with Merit Based Matching.

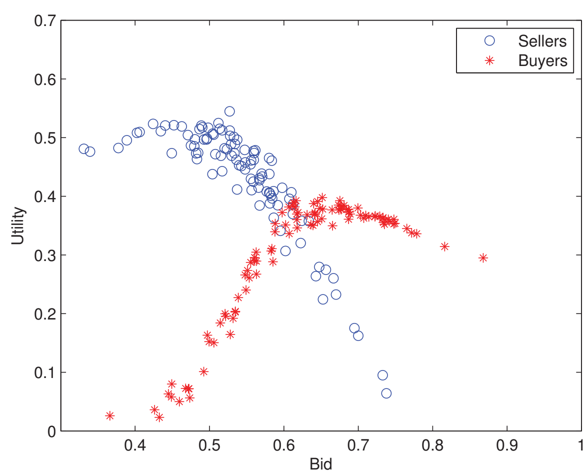
Step 0: Set the initial values of c, v, $\alpha > 1$, *and* $k = 0$. *Set initial values for* $\sigma_0 > 0$, $\mu_{b0} \in [c, v]$ *and* $\mu_{s0} \in [c, v]$. *For* $i, j = \{1, ..., n\}$ *draw samples of initial values of the ask and bid prices from uniform distributions over the interval* $[c, v]$, *i.e.*$p_{bi}(0) \sim U(c, v)$ *and* $p_{sj}(0) \sim U(c, v)$.



(a) Diagram of prices, Step 0



(b) Diagram of prices, Step 1



(c) Diagram of prices, Step 2

Figure 4–1: Evolution of prices in the algorithm 4.1, where $n = 100$, $m = 20$, $\sigma_0 = 0.3$, $\alpha = 1.1$, $c = 0$ and $v = 1$.

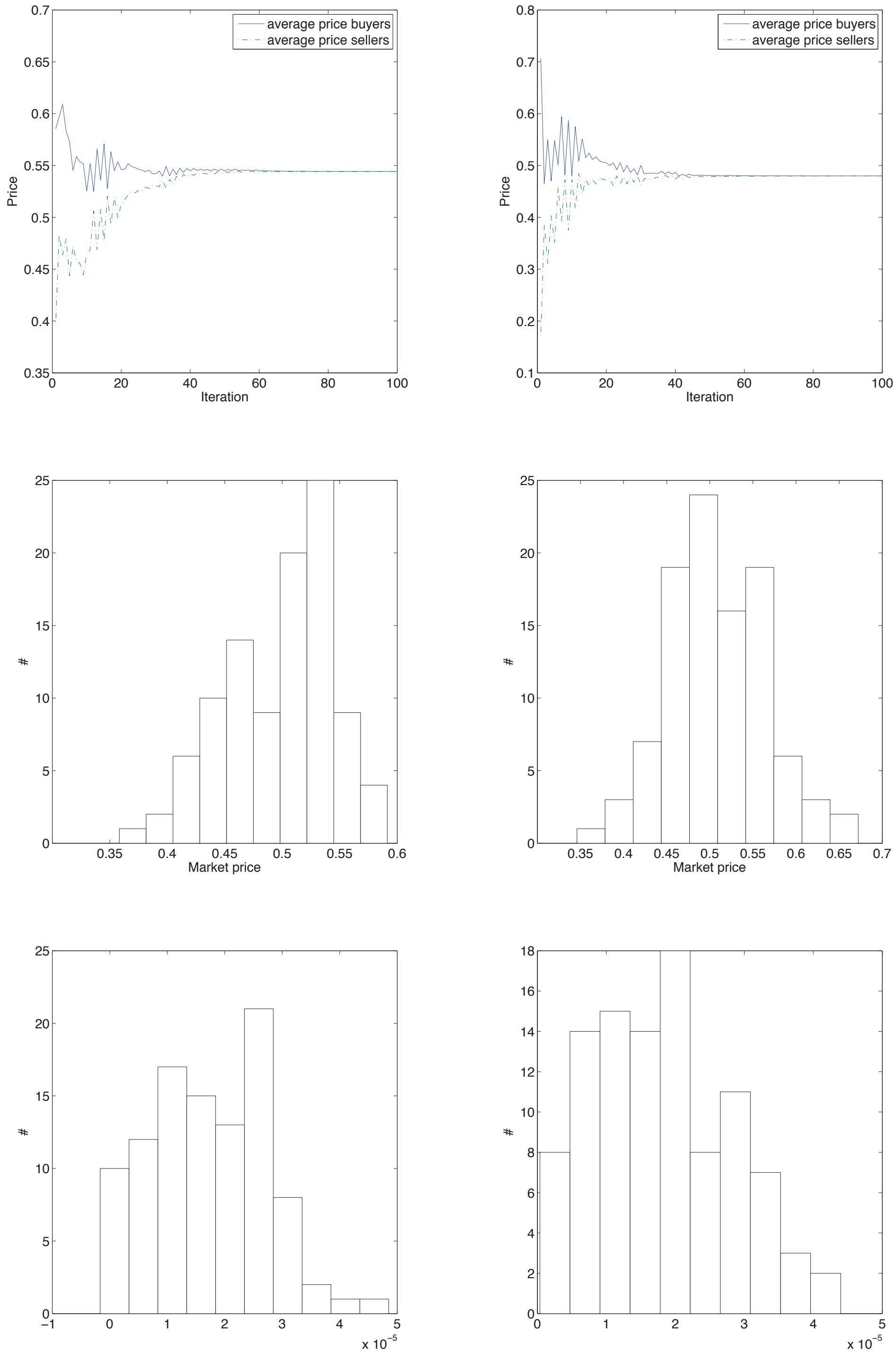


Figure 4–2: Convergence of the algorithm 4.1 after 100 iterations. In the left column for the utility functions (4.51) and (4.52), and in the right column for utility functions (4.53) and (4.54). In each case, the first row shows an example of convergence to an equilibrium. The second and the third rows show empirical frequencies of 100 trials for convergence to different market equilibria and the spread between the average bid and ask prices $(p_b - p_s)$, respectively. The parameters of the algorithm are chosen as $n = 100$, $m = 20$, $\sigma_0 = 0.3$, $\alpha = 1.1$, $c = 0$ and $v = 1$.

Step 1: Set $i = 1$, indicating that the utility function is calculated for buyer i.

Step 2: Calculate r, rank of buyer i among the buyers (buyer with the highest bid has rank 1). Find j the seller that has rank r among the sellers (seller with the lowest ask has rank 1).

Step 3: Use utility functions to calculate u_{bi}, and u_{sj} of the same rank.

Step 4: $i = i + 1$, go to Step 2 if $i < n + 1$.

Step 5: Update the price generator densities for the buyers and the sellers, as follows. First determine the indices i^* and j^* of the buyer and seller, respectively, who achieve the highest utility values in the current round of the game. Then set : $\mu_{bk} = p_{bi^*}$, and$\mu_{sk} = p_{sj^*}$.

Step 6: Evolve the price of each buyer and seller according to $p_{bi}(k+1) \sim N(\mu_{bk}, \sigma_k)$, $p_{sj}(k+1) \sim N(\mu_{sk}, \sigma_k)$; $i, j \in \{1, ..., n\}$. If $p_{bi} < c$, then $p_{bi} = c$. If $p_{bi} > v$, then $p_{bi} = v$. If $p_{sj} < c$, then $p_{sj} = c$. If $p_{sj} > v$, then $p_{sj} = v$.

Step 7: Contract the variance of the averages of the price generator densities for the buyers and the sellers: $\sigma_{k+1} = \sigma_k/\alpha$.

Step 8: Verify the algorithm's stopping condition. If $|\bar{p}_b(k) - \bar{p}_s(k)| > \epsilon$, then set $k = k + 1$, and go to Step 1, or else exit the algorithm.

Figure 4–3 represents the evolution of the average prices of the populations of buyers and sellers to a Nash equilibrium during the game. For this test the parameters of Algorithm 4.2 are chosen as follows : $n = 10$, $m = 20$, $v = 1$, $c = 0$, $\sigma_0 = 0.1$. For the left column $\alpha = 1.1$ and the test is ran for 100 iterations. For the right column , $\alpha = 1.01$, and the test is ran for 1000 iterations. The first row shows an example of convergence to an equilibrium. A comparison of the statistics of $p_b - p_s$, shown in the third row of Figure 4–3, indicates that a slower descent of variances (a lower α) combined with more iterations can lead to more convergence towards a Nash equilibrium. However, more tendency for convergence to one of the two extremes of the interval $[0, 1]$ is also observed in this case.

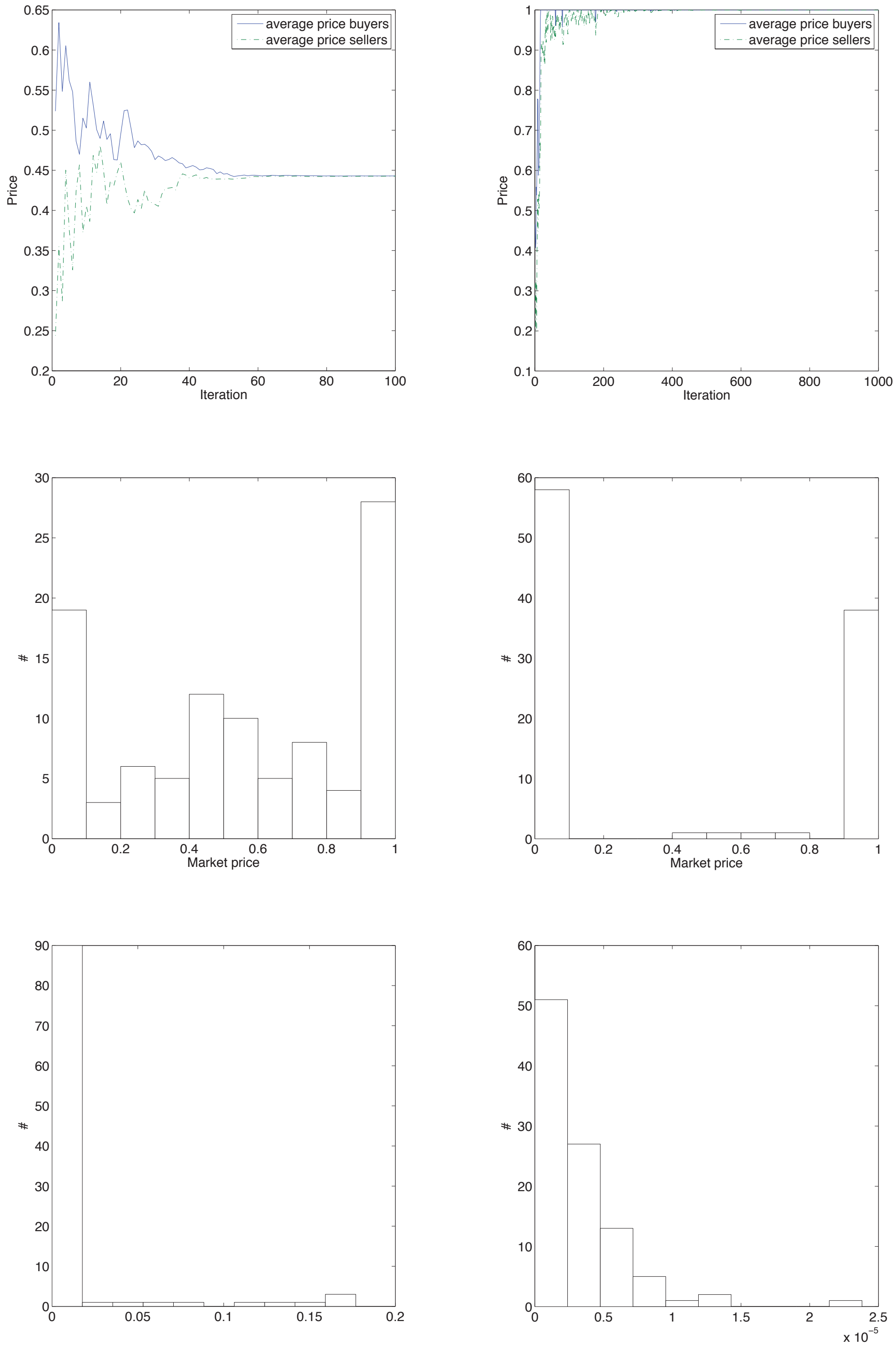


Figure 4–3: Convergence of the algorithm 4.2 while using utility functions (4.51) and (4.52). The parameters of the algorithm are chosen as $n = 10$, $v = 1$, $c = 0$, $\sigma_0 = 0.1$. For the left column, $\alpha = 1.1$, and the test is run for 100 iterations. For the right column, $\alpha = 1.01$, and the test is run for 1000 iterations. The first row shows an example of convergence to an equilibrium. The second and the third rows show empirical frequencies of 100 trials for convergence to different market equilibria, and the spread between the average bid and ask prices $(p_b - p_s)$, respectively.

4.6 Summary

An evolutionary algorithm is proposed for learning in double auction markets where the buyers and sellers follow the best member of their populations from the previous round of the game and mutate their bids by a diminishing Gaussian distribution.

The existence of a sequence of variances is proved that guarantees the convergence of the stochastic learning algorithm for the risk neutral and risk averse players to a Nash equilibrium.

Simulations show convergence when the sequence of variances is obtained by a geometric series when the buyers and sellers are matched randomly (Algorithm 4.1). Convergence does not happen as fast if buyers and sellers are matched based on their merits (Algorithm 4.2).

CHAPTER 5
A Random Search Algorithm for Learning in Double Auction Markets

As in Chapter 4, it is also of interest to find whether learning in the repeated game of double auction market can converge to an equilibrium.

To recapitulate, in a double auction market, buyers and sellers bid simultaneously. A transaction is possible if a buyer is matched with a seller that is offering a lower price than a buyer's bidding price. In such a game the utility functions will exhibit discontinuities.

In Chapter 5, convergence for an algorithm was analyzed where the buyers and the sellers follow the best bidders of their group. In this chapter a Random Search Algorithm (RSA) for two populations of players (i.e. buyers and sellers) is presented. The algorithm mimics the behaviour of buyers and sellers in a real market where individual buyers try to know what is the lowest limit that they can bid for buying and individual sellers try to know what is the upper limit that they can still sell. The speed of convergence of the algorithm is compared to the speed of convergence of the Genetic Algorithm proposed in [91].

5.1 Problem Statement

It is assumed that a double auction market can be modelled as in [91]: There are $n > 0$ buyers and the same number of sellers. In any round of the game any buyer will be matched with m random sellers. A transaction will take place, benefiting both the buyer and the seller only if the price bid by the buyer exceeds the price offered by the seller.

If $c \in [0, 1]$ is the cost of the production and $v \in [0, 1]$ represents the value of a good for the buyers, and under the assumption that a buyer and a seller will benefit from their transaction equivalently, the utility functions, u_{bi} and u_{sj}, of buyer i and seller j can be given by the formula below

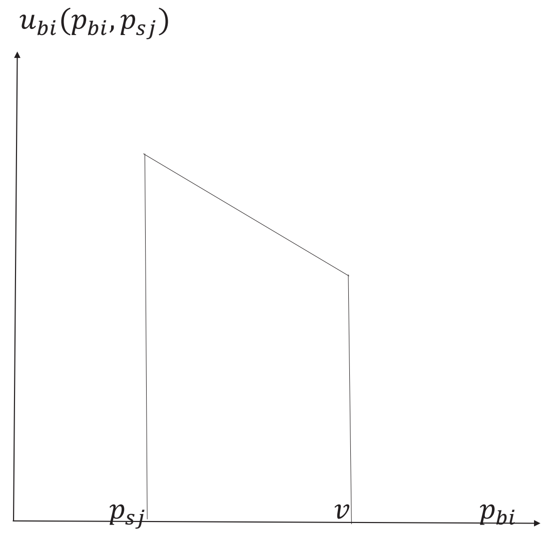


Figure 5–1: Discontinuity of utility function of buyer i

$$u_{bi}(p_{bi}, p_{sj}) = \begin{cases} v - \frac{p_{bi}+p_{sj}}{2} & \text{if } p_{bi} \in [p_{sj}, v] \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

$$u_{sj}(p_{bi}, p_{sj}) = \begin{cases} \frac{p_{bi}+p_{sj}}{2} - c & \text{if } p_{sj} \in [c, p_{bi}] \\ 0 & \text{otherwise} \end{cases}$$

in which p_{bi} and p_{sj} denote the prices of buyer i and seller j, respectively. Figures 5–1 and 5–2 illustrate the above utility functions. It can be seen that both utility functions are discontinuous and that the best price bid of a buyer and the best price asked by a seller are both very close to the point of discontinuity.

If all buyers and sellers bid the same price this game is in a Nash equilibrium (see Proposition 4.1). In [91] the agents learn to do better in each iteration by genetic reproduction from the successful agents of their appropriate group. Instead of using a genetic mutation to produce a candidate bid, the Random Search Algorithm explained in this chapter simulates a behaviour which uses the estimates for the best bids for buyers and sellers where they do not make assumptions about the probability distributions of each other in the next round of

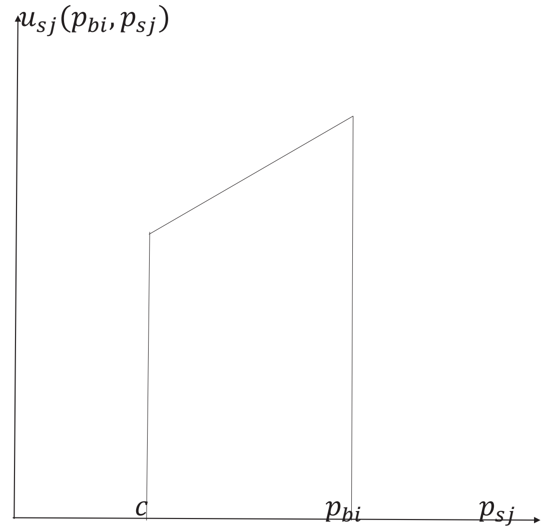


Figure 5–2: Discontinuity of utility function of seller j

the game. In addition, a vanishing Gaussian mutation is added that lets each population perform a random search against their counterparts, a behaviour that is more similar to actions in the market than the genetic learning. Hereafter, the problem is to know if this learning algorithm for the double auction market converges to a Nash equilibrium of the game, and if so how the speed of convergence is compared to the well-cited Genetic Algorithm for the same application [91].

5.2 Evolutionary Random Search Algorithm for Learning in Double Auction Markets

The evolutionary iterative algorithm for learning in double auction markets developed here belongs to the general class of random search algorithms. The underlying idea of the algorithm stems from a reasonable assumption that in double auction markets both the buyers and the sellers know about other buyers and sellers being either successful of unsuccessful in the preceding round of the game.

The following notation is adopted :

– n, the cardinality of the populations of buyers and sellers.

– k, the index of the current round of the game, ($k \in \mathbb{N}$).

– $c \in [0, 1]$, the cost of production for sellers.

– $v \in [0, 1]$, the value of the product for buyers.

– $p_{bi}(k) \in [c, v]$, the maximal price at which buyer i is willing to buy in round k.

– $p_{sj}(k) \in [c, v]$, the minimal price at which seller j is willing to sell in round k.

– $\bar{p}_b(k) \in [c, v]$, the average of the buyers' bid prices in round k, i.e., $\bar{p}_b(k) = \sum_{i=1}^{n} p_{bi}(k)/n$.

– $\bar{p}_s(k) \in [c, v]$, the average of the sellers' ask prices in round k, i.e., $\bar{p}_s(k) = \sum_{j=1}^{n} p_{sj}(k)/n$.

– $avsb$, the average price used by the successful buyers in a current round of the game.

– $avusb$, the average price used by the unsuccessful buyers in a current round of the game.

– $avss$, the average price used by the successful sellers in a current round of the game.

– $avuss$, the average price used by the unsuccessful sellers in a current round of the game.

– $borb$, a value representing a "boundary price" for buyers, as it evolves in the iterations of the algorithm, that separates the populations of successful and unsuccessful buyers.

– $bors$, a value representing a "boundary price" for sellers, as it evolves in the iterations of the algorithm, that separates the populations of successful and unsuccessful sellers.

– m, the number of buyers (sellers) that any seller (buyer) meets in any round of the game.

– $T_r(k)$, the proportion of how many times (out of m) a buyer or a seller must succeed in order to be labelled successful in round k of the evolution, $T_r(k) \in]0, 1[$.

– $\alpha > 1$, a shrinking factor for the variance of the randomizer function used in the generation of the bid prices.

– $\beta \in (0, .2)$, a margin distance between the means of the random generator functions and the corresponding borders of successful and unsuccessful buyers' bids or sellers' ask prices.

– $N(\mu, \sigma)$, the normal distribution with mean μ and variance σ.

– $\sigma_k > 0$, the variance of the random generator function in round k of the game.

– $\mu_{bk} \in [c, v]$ and $\mu_{sk} \in [c, v]$, the means for the random generator functions for buyers and sellers, respectively.

– m_{count}, a counter that controls the number of interactions of players before an update of prices.

– suc_{bi} and suc_{sj}, counters of successful transactions for buyer i and seller j, respectively; $i, j = \{1, \cdots, n\}$.

– s_{bi} and s_{sj}, flags to check if buyer i or seller j is marked as successful after meeting m counterparts.

– $\epsilon \in (0, 1)$, algorithm termination threshold.

Before stating the steps of the algorithm, it is helpful to explain the meaning behind them. The values of the algorithm parameters and the initial values of the buyers and sellers prices are selected in Steps 0 and 1. The latter are the initial maximal prices at which the buyers are willing to buy and the initial minimal prices at which the sellers are willing to sell. Steps 2 - 5 constitute a loop in which each buyer meets m sellers in the current round of the algorithm. Each buyer or seller who succeeded in at least $T_r(k).m$ transactions is then labelled as successful. At the exit of the loop in Steps 2 - 5, the average prices of each of the so categorized four groups of buyers and sellers are then computed (Step 6) yielding the average prices for successful as well as unsuccessful buyers and sellers. The averaging then account of the updated prices of successful buyers and sellers. The averages are next used to evolve the boundary prices $borb$ and $bors$ that separate the populations of successful and unsuccessful buyers and sellers. The latter are employed to shape the probability distributions of the buyers' and the sellers' maximal bidding prices and minimal asking prices, see Steps 7-8. In Step 9, new prices are drawn for all the buyers and sellers. Again, these prices represent the current maximal prices at which the buyers are willing to buy and the minimal prices at which the sellers are willing to sell. The variance of both probability distributions are shrunk by a factor $1/\alpha$ for the next round of the game. It is the evolution of the boundary prices separating the successful and unsuccessful players that drives the evolution of the randomizing probability distributions for players in the market.

These boundaries approach one another as players learn about the market whose behaviour is tightly related to the ensemble of their utility functions. The algorithm is exited if the prices of the buyers and sellers are sufficiently close to each other (close to the equilibrium of the game).

Clearly, the information structure in this game is as follows: the players know their own utility functions, their own price and the current price of their opponents in the market game, and they can also "estimate" the current average prices employed in their own group (of either buyers or sellers). Algorithm 5.1 can thus be viewed as an algorithm for the simulation of the evolution of a double auction market in which players can learn progressively how to maximize their gains, respective to the ensemble of the utility functions of the players. The evolution terminates when a market equilibrium is reached. The value of the equilibrium clearly depends on the rate at which the variances of the randomizing probability distributions for the prices of the players approach zero. The evolution of these variances is meant to represent the readiness of the players to adhere to the market trend.

Algorithm 5.1 The Random Search Algorithm. *Step 0: Set the initial values of* m, $T_r(0)$, $\alpha > 1$, $\beta \in [0, .2)$, *and* $k = 0$. *Set initial values for* $\sigma_0 > 0$, $\mu_{b0} \in [c, v]$ *and* $\mu_{s0} \in [c, v]$. *For* $i, j = \{0, ..., n\}$ *draw samples of initial values of the ask and bid prices from uniform distributions over the interval* $[c, v]$, *i.e.*$p_{bi}(0) \sim U(c, v)$ *and* $p_{sj}(0) \sim U(c, v)$.

Step 1: Set $m_{\text{count}} = 0$ *and* $suc_{bi} = suc_{sj} = 0$, $i, j \in \{1, ..., n\}$.

Step 2: Draw an integer $i^* \in \{1, 2, ..., n\}$ *from a uniform distribution (i.e.* $Pr(i^*) = 1/n$*). Draw an integer* $j^* \in \{1, 2, ..., n\}$ *from a uniform distribution (i.e.* $Pr(j^*) = 1/n$*).*

Step 3: For each $i \in \{1, ..., n\}$ *if (* $p_{bi^*} > p_{sj^*}$ *) then increment the counters of successful buyer and successful seller as follows:* $suc_{bi^*} = suc_{bi^*} + 1$, $suc_{sj^*} = suc_{sj^*} + 1$.

Step 4: Increment counter $m_{\text{count}} = m_{\text{count}} + 1$.

Step 5: if $m_{\text{count}} < m$, *go to Step 2.*

Step 6: Label buyers as successful and unsuccessful: if $suc_{bi} > T_r(k).m$, *then set* $s_{bi} = 1$ *else set* $s_{bi} = 0$. *Label sellers as successful and unsuccessful: if* $suc_{sj} > T_r(k).m$, *then set* $s_{sj} = 1$

else set $s_{sj} = 0$.

Step 7: Using the labels set in Step 6, calculate the average price used by the successful buyers, avsb, the average price used by the unsuccessful buyers, avusb, the average price used by the successful sellers, avss, and the average price used by the unsuccessful sellers: avuss.

Step 8: Set $borb = (avsb + avusb)/2$, $bors = (avss + avuss)/2$.

Step 9: Update the averages of the price generator densities for the buyers and the sellers, respectively: $\mu_{bk} = borb - \beta\sigma_k$, *and* $\mu_{sk} = bors + \beta\sigma_k$. *If* $\mu_{bk} \notin [c, v]$ *then* μ_{bk} *is reset to* $\mu_{bk} = c$ *or* $\mu_{bk} = v$ *depending on whether* $\mu_{bk} < c$ *or* $\mu_{bk} > v$. *Similar truncation is also performed for* μ_{sk}.

Step 10: Evolve the price of each buyer and seller according to $p_{bi}(k+1) \sim N(\mu_{bk}, \sigma_k)$, $p_{sj}(k+1) \sim N(\mu_{sk}, \sigma_k)$; $i, j \in \{1, ..., n\}$. *If* $p_{bi} < c$, *then* $p_{bi} = c$. *If* $p_{bi} > v$, *then* $p_{bi} = v$. *If* $p_{sj} < c$, *then* $p_{sj} = c$. *If* $p_{sj} > v$, *then* $p_{sj} = v$.

Step 11: Contract the variance of the price generator densities for the buyers and the sellers: $\sigma_{k+1} = \sigma_k/\alpha$.

Step 12: Verify the algorithm's stopping condition. If $|\bar{p}_b(k) - \bar{p}_s(k)| > \epsilon$, *update* $T_r(k)$, *e.g.,* $T_r(k+1) = T_r(k)$, *then set* $k = k+1$, *and go to Step 1, else exit the algorithm.*

5.3 Proof of Convergence

It can be seen from Steps 8-9 and Step 11 that the probability distributions of the prices are modified in a way which favors the averages to move towards meeting one another (producing an equilibrium of the game).

Remark 5.1 *The notation* $N(x; \mu, \sigma)$ *is used for the normal probability distribution function with mean* μ *and variance* σ, *and the notation* $\Phi(x)$ *is used for the cumulative distribution function of this distribution,* $\Phi(x) = \int_{-\infty}^{x} N(\zeta; \mu, \sigma)d\zeta$.

Proposition 5.1 *Assume that for* n, *the cardinality of the populations of the players and* m, *the number of their counterparts in each round of the game, the probability of success for buyer* i *is denoted by* $Pr_{nm}(Success\ for\ buyer\ i|p_{bi})$ *and the probability of success for the seller* j *is denoted by* $Pr_{nm}(Success\ for\ seller\ j|p_{sj})$. *Furthermore, let us define these two* *limits*

$$\lim_{n,m\to\infty} Pr_{nm}(Success\ for\ buyer\ i|p_{bi}) \triangleq Pr(Success\ for\ buyer\ i|p_{bi}) \tag{5.2}$$

$$\lim_{n,m\to\infty} Pr_{nm}(Success\ for\ seller\ j|p_{sj}) \triangleq Pr(Success\ for\ seller\ j|p_{sj}). \tag{5.3}$$

If for every $T_r(k) \in]0,1[$, x_{ob} and x_{os} are chosen such that $\int_{-\infty}^{x_{ob}} N(x;\mu_{sk},\sigma_k)dx = T_r(k)$, and $\int_{x_{os}}^{\infty} N(x;\mu_{bk},\sigma_k)dx = T_r(k)$, then

$$Pr(Success\ for\ buyer\ i|p_{bi}) = \begin{cases} 0 & if\ p_{bi} \le x_{ob} \\ 1 & otherwise \end{cases} \tag{5.4}$$

$$Pr(Success\ for\ seller\ j|p_{sj}) = \begin{cases} 0 & if\ p_{sj} > x_{os} \\ 1 & otherwise. \end{cases} \tag{5.5}$$

Proof. To prove (5.4), starting from (5.2),

$$\lim_{n,m\to\infty} Pr_{nm}(\text{Success for buyer } i|p_{bi}) \triangleq Pr(\text{Success for buyer } i|p_{bi}), \tag{5.6}$$

and by using Step 6 of the Algorithm,

$$= \lim_{n,m\to\infty} Pr(\frac{\text{Number of times out of } m' \text{ that } p_{bi} > p_{sj^*}}{m'} > T_r(k)). \tag{5.7}$$

The index j^* can indicate a different seller in each of the m' trials, which is proportional to m because matching happens with equal probability. Since the buyer i, who bids p_{bi} is matched with probability $1/n$ with a random seller who bids p_{sj^*}, and p_{sj^*} is drawn from a Normal distribution in the Step 10, for $p_{bi} = x_{ob}$,

$$\lim_{n,m\to\infty} \frac{\text{Number of times out of } m' \text{ that } p_{bi} \ge p_{sj^*}}{m'} \tag{5.8}$$

$$= Pr(p_{sj^*} \le x_{ob}) \tag{5.9}$$

$$= \int_{-\infty}^{x_{ob}} N(x;\mu_{sk},\sigma_k)dx = T_r(k), \tag{5.10}$$

and (5.4) is concluded. Proof of (5.5) is similar.

Corollary 5.1 *x_{ob} and x_{os}, which are calculated in the proposition above for $T_r(k)$, satisfy these two relations:*

$$x_{ob} - \mu_{bk} = \mu_{sk} - x_{os} \tag{5.11}$$

and

$$\Phi(x_{ob} - \mu_{bk}) = 1 - \Phi(x_{os} - \mu_{sk}) \tag{5.12}$$

Proof. Trivial.

Theorem 5.1 *For every $\epsilon > 0$, there exist $n_0, m_0 \in \mathbb{N}$ such that if $n > n_0$ (the cardinality of the populations of the players) and $m > m_0$ (the number of counterparts met in each round of the game) then there exist a contraction factor $\alpha > 1$ and a sequence of thresholds for labelling a buyer or a seller as successful, i.e., $\exists\{T_r(k)\}_{k=0}^{\infty}$ such that $\lim_{k\to\infty} Pr(|\mu_{bk} - \mu_{sk}| > \epsilon) = 0$, which implies that a Nash equilibrium is achieved.*

Proof. At time k, choosing a $T_r(k) \in]0,1[$, x_{ob} and x_{os} are calculated from (5.1) and the proof starts with calculating the average bids of successful and unsuccessful buyers and sellers, $\overline{SB}$, $\overline{USB}$, $\overline{SS}$ and $\overline{USS}$.

First, average price of bids for the successful buyers is calculated

$$\begin{aligned}
\overline{SB} &\triangleq \int_{-\infty}^{\infty} xPr(p_{bi} = x| \text{ buyer } i \text{ is successful})dx \\
&= \frac{\int_{-\infty}^{\infty} xPr(\text{ buyer } i \text{ is successful}|p_{bi} = x)Pr(x)dx}{Pr(\text{ buyer } i \text{ is successful})}.
\end{aligned}$$

Using Proposition 5.1, the denominator is calculated as

$$D_{SB} \triangleq Pr(\text{ buyer } i \text{ is successful}) = \int_{x_{ob}}^{\infty} N(x; \mu_{sk}, \sigma_k)dx = 1 - \Phi(\frac{x_{ob} - \mu_{sk}}{\sigma_k}) \tag{5.13}$$

where x_{ob} is defined in Proposition 5.1, and the numerator is calculated as

$$\begin{aligned}
N_{SB} &\triangleq \int_{-\infty}^{\infty} xPr(\text{buyer } i \text{ is successful}|p_{bi} = x)Pr(x)dx \\
&= \int_{x_{ob}}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_k} e^{\frac{-(x-\mu_{bk})^2}{2\sigma_k^2}} dx.
\end{aligned} \tag{5.14}$$

By the change of variable $x - \mu_{bk} = \zeta$

$$\begin{aligned}
N_{SB} &= \int_{x_{ob}-\mu_{bk}}^{\infty} \frac{\zeta + \mu_{bk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \int_{x_{ob}-\mu_{bk}}^{\infty} \frac{\zeta}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta + \int_{x_{ob}-\mu_{bk}}^{\infty} \frac{\mu_{bk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \frac{\sigma_k}{\sqrt{2\pi}} \int_{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}}^{\infty} e^{-\alpha} d\alpha + \mu_{bk}(1 - \Phi(\frac{x_{ob} - \mu_{bk}}{\sigma_k})) \\
&= \frac{\sigma_k}{\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} + \mu_{bk}(1 - \Phi(\frac{x_{ob} - \mu_{bk}}{\sigma_k}))
\end{aligned}$$

Hence, $\overline{SB}$ is calculated as

$$\overline{SB} = \frac{N_{SB}}{D_{SB}} = \frac{\sigma_k}{\sqrt{2\pi}(1 - \Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k}))} e^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} + \mu_{bk}. \tag{5.15}$$

Then, the average price of bids for the unsuccessful buyers is calculated as

$$\begin{aligned}
\overline{USB} &\triangleq \int_{-\infty}^{\infty} xPr(p_{bi} = x| \text{ buyer } i \text{ is unsuccessful})dx \\
&= \frac{\int_{-\infty}^{\infty} xPr(\text{ buyer i is unsuccessful}|p_{bi} = x)Pr(x)dx}{Pr(\text{ buyer i is unsuccessful})}.
\end{aligned}$$

Using Proposition 5.1, the denominator is calculated as

$$D_{SB} \triangleq Pr(\text{ buyer } i \text{ is unsuccessful }) = \int_{\infty}^{x_{ob}} N(x; \mu_{sk}, \sigma_k)dx = \Phi(\frac{x_{ob} - \mu_{sk}}{\sigma_k}) \tag{5.16}$$

and the numerator is calculated as

$$N_{USB} \triangleq \int_{-\infty}^{\infty} xPr(\text{ buyer } i \text{ is unsuccessful } |p_{bi} = x)Pr(x)dx \tag{5.17}$$

$$= \int_{-\infty}^{x_{ob}} \frac{x}{\sqrt{2\pi}\sigma_k} e^{\frac{-(x-\mu_{bk})^2}{2\sigma_k^2}} dx. \tag{5.18}$$

By the change of variable $x - \mu_{bk} = \zeta$

$$
\begin{aligned}
N_{USB} &= \int_{-\infty}^{x_{ob}-\mu_{bk}} \frac{\zeta+\mu_{bk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \int_{-\infty}^{x_{ob}-\mu_{bk}} \frac{\zeta}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta + \int_{-\infty}^{x_{ob}-\mu_{bk}} \frac{\mu_{bk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \frac{-\sigma_k}{\sqrt{2\pi}} \int_{-\infty}^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} e^{-\alpha} d\alpha + \mu_{bk}\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k}) \\
&= \frac{-\sigma_k}{\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} + \mu_{bk}\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k}).
\end{aligned}
$$

Hence, $\overline{USB}$ is calculated as

$$
\overline{USB} = \frac{N_{USB}}{D_{USB}} = \frac{-\sigma_k}{\sqrt{2\pi}\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} e^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} + \mu_{bk}. \tag{5.19}
$$

Finally, the $\overline{borb}$ is defined as

$$
\overline{borb} = \frac{\overline{SB}+\overline{USB}}{2} \tag{5.20}
$$

$$
\overline{borb} = \mu_{bk} + \frac{\sigma_k}{2\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} \left(\frac{1}{1-\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} - \frac{1}{\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})}\right) \tag{5.21}
$$

the border for successful and unsuccessful sellers, $\overline{bors}$, is also calculated

$$
\begin{aligned}
\overline{SS} &\triangleq \int_{-\infty}^{\infty} xPr(p_{sj}=x|\text{ seller j is successful})dx \\
&= \frac{\int_{-\infty}^{\infty} xPr(\text{ seller j is successful}|p_{sj}=x)Pr(x)dx}{Pr(\text{ seller j is successful})}.
\end{aligned}
$$

The denominator is calculated as

$$
Pr(\text{ seller } j \text{ is successful}) = \int_{x_{ob}}^{\infty} N(x,\mu_{sk},\sigma_k)dx = \Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k}) \tag{5.22}
$$

and the numerator is calculated as

$$
\int_{-\infty}^{\infty} xPr(\text{ seller } j \text{ is successful}|p_{sj}=x)Pr(x)dx \tag{5.23}
$$

$$
= \int_{-\infty}^{x_{os}} \frac{x}{\sqrt{2\pi}\sigma_k} e^{\frac{-(x-\mu_{sk})^2}{2\sigma_k^2}} dx. \tag{5.24}
$$

By introducing the variable $\zeta \triangleq x - \mu_{sk}$

$$\begin{aligned}
N_{SS} &= \int_{-\infty}^{x_{os}-\mu_{sk}} \frac{\zeta + \mu_{sk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \int_{-\infty}^{x_{os}-\mu_{sk}} \frac{\zeta}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta + \int_{-\infty}^{x_{os}-\mu_{sk}} \frac{\mu_{bk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \frac{\sigma_k}{\sqrt{2\pi}} \int_{-\infty}^{\frac{(x_{ob}-\mu_{sk})^2}{2\sigma_k^2}} e^{-\alpha} d\alpha + \mu_{sk}\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k}) \\
&= \frac{-\sigma_k}{\sqrt{2\pi}} e^{\frac{-(x_{0s}-\mu_{sk})^2}{2\sigma_k^2}} + \mu_{sk}\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k}).
\end{aligned}$$

Hence, $\overline{SS}$ is calculated as

$$\overline{SS} = \frac{N_{SS}}{D_{SS}} = \frac{-\sigma_k}{\sqrt{2\pi}\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k})} e^{\frac{-(x_{os}-\mu_{sk})^2}{2\sigma_k^2}} + \mu_{sk}. \tag{5.25}$$

Finally, the average price of bids for the unsuccessful sellers is calculated

$$\begin{aligned}
\overline{USS} &\triangleq \int_{-\infty}^{\infty} xPr(p_{sj} = x| \text{ seller j is unsuccessful})dx \\
&= \frac{\int_{-\infty}^{\infty} xPr(\text{ seller j is unsuccessful}|p_{sj} = x)Pr(x)dx}{Pr(\text{ seller j is unsuccessful})}.
\end{aligned}$$

Using Proposition 5.1, the denominator is calculated as

$$D_{USS} \triangleq Pr(\text{ seller } j \text{ is unsuccessful}) = \int_{x_{os}}^{\infty} N(x, \mu_{sk}, \sigma_k)dx = 1 - \Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k}) \tag{5.26}$$

and the numerator is calculated as

$$N_{USS} \triangleq \int_{-\infty}^{\infty} xPr(\text{ seller } j \text{ is unsuccessful}|p_{sj} = x)Pr(x)dx \tag{5.27}$$

$$= \int_{x_{os}}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_k} e^{\frac{-(x-\mu_{sk})^2}{2\sigma_k^2}} dx. \tag{5.28}$$

By introducing variable $\zeta = x - \mu_{sk}$

$$
\begin{aligned}
N_{USS} &= \int_{x_{os}-\mu_{sk}}^{\infty} \frac{\zeta+\mu_{sk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \int_{x_{os}-\mu_{sk}}^{\infty} \frac{\zeta}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta + \int_{x_{os}-\mu_{sk}}^{\infty} \frac{\mu_{sk}}{\sqrt{2\pi}\sigma_k} e^{\frac{-\zeta^2}{2\sigma_k^2}} d\zeta \\
&= \frac{\sigma_k}{\sqrt{2\pi}} \int_{\frac{-(x_{os}-\mu_{sk})^2}{2\sigma_k^2}}^{\infty} e^{-\alpha} d\alpha + \mu_{sk}(1-\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k})) \\
&= \frac{\sigma_k}{\sqrt{2\pi}} e^{\frac{-(x_{os}-\mu_{sk})^2}{2\sigma_k^2}} + \mu_{sk}(1-\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k})).
\end{aligned}
$$

Hence, $\overline{USS}$ is calculated as

$$
\overline{USS} = \frac{N_{USS}}{D_{USS}} = \frac{\sigma_k}{\sqrt{2\pi}(1-\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k}))} e^{\frac{-(x_{os}-\mu_{sk})^2}{2\sigma_k^2}} + \mu_{sk}. \tag{5.29}
$$

Then, $\overline{bors}$ is defined as

$$
\overline{bors} \triangleq \frac{\overline{SS}+\overline{USS}}{2} \tag{5.30}
$$

$$
\overline{bors} = \mu_{sk} + \frac{\sigma_k}{2\sqrt{2\pi}} e^{\frac{-(x_{os}-\mu_{sk})^2}{2\sigma_k^2}} (\frac{1}{1-\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k})} - \frac{1}{\Phi(\frac{x_{os}-\mu_{sk}}{\sigma_k})}). \tag{5.31}
$$

Recall from Corollary 5.1 that $x_{ob}-\mu_{bk} = \mu_{sk}-x_{os}$, and $\Phi(x_{ob}-\mu_{bk}) = 1-\Phi(x_{os}-\mu_{sk})$. Hence (5.31) can be written as

$$
\overline{bors} = \mu_{sk} + \frac{\sigma_k}{2\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{2\sigma_k^2}} (\frac{1}{\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} - \frac{1}{1-\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})}). \tag{5.32}
$$

Then, the difference "$\overline{borb}-\overline{bors}$" which indicates contraction is calculated as

$$
\overline{borb}-\overline{bors} = \mu_{bk}-\mu_{sk} + \frac{\sigma_k}{\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{\sigma_k^2}} (\frac{1}{1-\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} - \frac{1}{\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})}). \tag{5.33}
$$

Let $d_{k+1} \triangleq \mu_{b(k+1)} - \mu_{s(k+1)} = (borb+\beta\sigma_k) - (bors-\beta\sigma_k)$, and $\beta \in [0, 0.2)$ be an arbitrary chosen scalar. For $\beta = 0$, it follows that

$$
d_{k+1} \triangleq \mu_{b(k+1)} - \mu_{s(k+1)} \tag{5.34}
$$

since $borb$ and $bors$ are calculated in the Step 8, and $\overline{borb}$ and $\overline{bors}$ are calculated by the integrals.

Using Proposition 5.1, and all the calculations above used to find $\overline{borb}$ and $\overline{bors}$, it is concluded that for every $\epsilon_1, \epsilon_2 > 0$ there exists $n_0, m_0, k_0 \in N$ such that for every $n > n_0$, $m > m_0$, $k > k_0$ ($\sigma_{k+1} = \sigma_k/\alpha$), if

$$Pr(\mu_{bk}, \mu_{sk} \in]c, v[) = 1 \tag{5.35}$$

then

$$Pr(|(borb - bors) - (\overline{borb} - \overline{bors})| > \epsilon_1) < \epsilon_2 \tag{5.36}$$

and by using (5.33) and (5.34)

$$Pr(|d_{k+1} - d_k - \frac{\sigma_k}{\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{\sigma_k^2}} (\frac{1}{1 - \Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} - \frac{1}{\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})})| > \epsilon_1) < \epsilon_2. \tag{5.37}$$

If for every $k \in N$

$$\frac{1}{\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{\sigma_k^2}} (\frac{1}{1 - \Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} - \frac{1}{\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})}) = \Psi \tag{5.38}$$

then (5.37) reduces to

$$Pr(|d_{k+1} - d_k - \sigma_k \Psi| > \epsilon_1) < \epsilon_2 \tag{5.39}$$

(5.35) and (5.39) together result

$$Pr(\mu_{b(k+1)}, \mu_{s(k+1)} \in]c, v[) = 1 \tag{5.40}$$

and for the time $k' > k > k_0$,

$$Pr(|d_{k'} - d_k - \Psi \sum_{l=k}^{k'-1} \sigma_l| > (k' - k)\epsilon_1) < \epsilon_2 \tag{5.41}$$

and since $\sigma_{k+1} = \sigma_k/\alpha$,

$$Pr(|d_{k'} - d_k - \Psi\sigma_0 \sum_{l=k}^{k'-1} \frac{1}{\alpha^l}| > \epsilon_1') < \epsilon_2'. \tag{5.42}$$

where, $\epsilon_1' = (k' - k)\epsilon_1$ and $\epsilon_2' = \epsilon_2$. Next, we invoke that if

$$\Psi\sigma_0 \Sigma_{l=k}^{\infty} \frac{1}{\alpha^l} = -d_k, \tag{5.43}$$

the theorem is proven, because there exist $k'_0, n_0, m_0 \in N$ such that for every $\epsilon'_1, \epsilon'_2 > 0$ $k' > k'_0$, $n > n_0$, and $m > m_0$,

$$Pr(|d_{k'}| > \epsilon_1) < \epsilon_2, \tag{5.44}$$

where $d_{k'} = \mu_{bk'} - \mu_{sk'}$.

To this end, since $\alpha > 1$, (5.43) reduces to

$$\Psi = \frac{-d_k \alpha^{k-1}(\alpha - 1)}{\sigma_0}. \tag{5.45}$$

Therefore, if $\mu_{bk} \in]c, v[$, since we have (5.40), (5.45) is valid for every $k' > k$ (if not equation (5.36) could not be used), and because the left side of the equation (5.38) is equal to zero for $x_{ob} = \mu_{bk}$, the convergence is ensured if for every $k'' > k'_0$, there exist $\alpha > 1$ and $T_r(k'') \in]0, 1[$ such that x_{ob} calculated from Proposition 5.1 for such a $T_r(k'')$ satisfies

$$\frac{1}{\sqrt{2\pi}} e^{\frac{-(x_{ob}-\mu_{bk})^2}{\sigma_k^2}} \left(\frac{1}{1 - \Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})} - \frac{1}{\Phi(\frac{x_{ob}-\mu_{bk}}{\sigma_k})}\right) = \frac{-d_k \alpha^{k-1}(\alpha - 1)}{\sigma_0}. \tag{5.46}$$

If $\mu_{bk} = c$, by the definition of x_{os} in Proposition 5.1, there exists $\gamma_0 > 0$ such that for every $\gamma_0 > \gamma > 0$, there exists a $T_r(k) \in]0, 0.5[$, less than half of population, such that $x_{os} > c + \gamma$. Consequently, because of Step 8, there exists $\gamma' > 0$ such that $bors > c + \gamma'$, (since average of unsuccessful sellers is bigger than c), and because of step 9, $\mu_{sk} > c + \gamma'$ (if $\beta = 0$) . The same is true for $\mu_{bk} = v$, $\mu_{sk} = c$ and $\mu_{sk} = v$. Thus, at every round k of the game, there exists $T_r(k)$ such that $Pr(\mu_{b(k+1)} \in]c, v[) = 1$, $Pr(\mu_{s(k+1)} \in]c, v[) = 1$, and from time $k + 1$ the existence for a sequence $\{T_r(k)\}_{k=0}^{\infty}$ can be proven as above.

Consequently, there exists $\alpha > 1$ and $\{T_r(k)\}_{k=0}^{\infty}$ such that convergence to a Nash equilibrium is guaranteed.

5.4 Genetic Learning as a Competitor Algorithm

The proposed random search algorithm is compared to the genetic algorithm used in [91] to observe its learning procedure in a double auction market with the same utility functions as adopted here. The competitor algorithm is summarized in Algorithm 5.2. To seek a fair comparison the structure of genetic algorithm as well as all the assumptions are chosen as close as possible to [91].

The genetic algorithm of [91] is summarized below for easier comparison.

Algorithm 5.2 The Genetic Algorithm *[91] :*

Step 0: Draw samples of initial values of p_{bi} *and* p_{sj} $\forall i, \forall j$ *as in Step 0 of Algorithm 5.1.*

Step 1: Let every buyer meet m *random sellers (as in Steps 2-5 of Algorithm 5.1).*

Step 2: Rank buyers and sellers according to the values of a fitness function selected. The fitness function for a buyer is calculated as the sum of the values of his/her utility function as he/she meets with all the m *sellers. The fitness function for sellers are defined similarly except that summation is performed over all the buyers which they meet.*

Step 3: Retain buyers and sellers whose ranking exceeds a certain threshold.

Step 4: Using crossover function of the genetic algorithm, re-generate new values for any p_{bi} *and* p_{sj} *that were eliminated in Step 3.*

Step 5: Employing the mutation function of the genetic algorithm, update p_{bi} *and* p_{sj} $\forall i, \forall j$ *to new values.*

Step 6: Go to Step 1 if the stopping conditions (similar to those of Step 12 in Algorithm 5.1) are not satisfied.

5.5 Numerical Experiments

Convergence properties of the novel random algorithm and the genetic algorithm of [91] are compared by way of simulations.

It is assumed that the cardinality of the populations of buyers and sellers is $n = 100$. The initial prices bid by the buyers and those asked by the sellers are drawn from a uniform distribution function on the interval $[0, 1]$.

In the case of the random search algorithm, the algorithm parameters of Step 0 were chosen by trial and error and as such were not selected optimally to secure fastest convergence possible.

The genetic algorithm of [91] was used with a mutation coefficient $\mu = .001$ (probability of changing a single bit in the digital representation of a price), as was assumed in [91]. Players decide to change their bids after $m = 20$ mutual encounters of buyers and sellers. The total of 20 iterations of evolution of the algorithm were simulated. It was found that small changes in the structure and the value of parameters of the genetic algorithm did not lead to results which differ much from the ones presented.

The values of parameters in Algorithm 5.1 were chosen as follows: $\sigma_0 = 0.1$, $\alpha = 1.2$, $\beta = 0$, and $Tr_k(0) = 0.5$, for all k.

The numerical tests are once run for this pair of utility functions [91] :

$$u_{bi}(p_{bi}, p_{sj}) = \begin{cases} v - \frac{p_{bi}+p_{sj}}{2} & \text{if } p_{bi} \in [p_{sj}, v] \\ 0 & \text{otherwise} \end{cases} \tag{5.47}$$

$$u_{sj}(p_{bi}, p_{sj}) = \begin{cases} \frac{p_{bi}+p_{sj}}{2} - c & \text{if } p_{sj} \in [c, p_{bi}] \\ 0 & \text{otherwise} \end{cases} \tag{5.48}$$

and the second time for this pair of utility functions:

$$u_{bi}(p_{bi}, p_{sj}) = \begin{cases} v - (\frac{p_{bi}+p_{sj}}{2})^2 & \text{if } p_{bi} \in [p_{sj}, v] \\ 0 & \text{otherwise} \end{cases} \tag{5.49}$$

$$u_{sj}(p_{bi}, p_{sj}) = \begin{cases} (\frac{p_{bi}+p_{sj}}{2})^2 - c & \text{if } p_{sj} \in [c, p_{bi}] \\ 0 & \text{otherwise} \end{cases} \tag{5.50}$$

5.5.1 Speed of Convergence

Figure 5–3 compares the convergence of the two algorithms with parameters as specified above while using the utility functions of formulas (5.47) and (5.48) or (5.49) and (5.50).

Figure 5–3(a), shows the evolution of the average prices of the buyers and sellers in their convergence to an equilibrium of the markets when the genetic algorithm is employed and mutation is not permitted. It is seen that the buyers' and sellers' average prices converge to a constant which then becomes an equilibrium point (see [91]).

Figure 5–3(c), shows a sample of how average prices for buyers and sellers converge when the genetic algorithm is implemented with mutation (see [91]).

Figure 5–3(e), depicts the evolution of the average prices of buyers and sellers as generated by the novel algorithm. Clearly, the convergence rate is much faster than that achieved using the genetic algorithm.

A better comparison for the speeds of convergence of the genetic and random search algorithms can be found in the left columns of Figures 5–4 to 5–6 where the histogram of the average bid and ask spread, i.e. $p_b - p_s$ is shown after 5, 10, and 15 rounds of the game respectively. The gap $p_b - p_s$ is chosen as an indicator of the speed of convergence because $p_b = p_s$ means convergence to Nash equilibrium for the average.

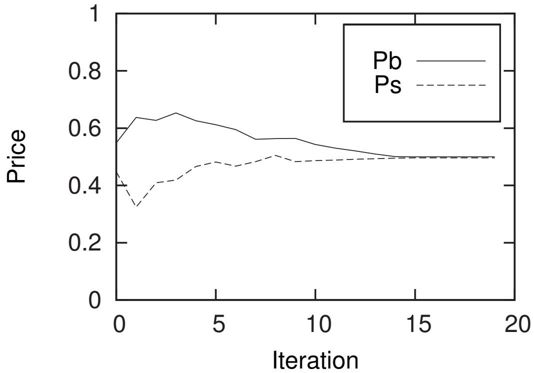
The new algorithm, Algorithm 5.1, is tested and compared to the genetic algorithm using a nonlinear pair of utility functions too. In the right column of Figure 5–3, the two algorithms are compared with parameters as specified above while using the utility functions of formulae (5.49) and (5.50).

Similarly, the right columns of Figures 5–4 to 5–6 show more details about the speed of convergence while using the utility functions of formulae (5.49) and (5.50).

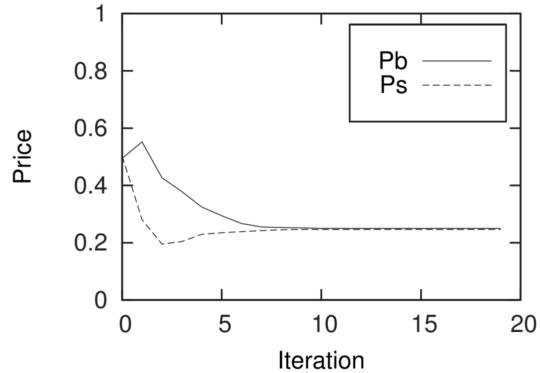
Figures 5–7 shows histograms of the market price equilibria as found by the genetic algorithm and the random search algorithm. It is seen that the probability distribution of the market equilibria achieved by the genetic algorithm is characterized by a much larger variance.

5.5.2 Small Populations of Buyers and Sellers

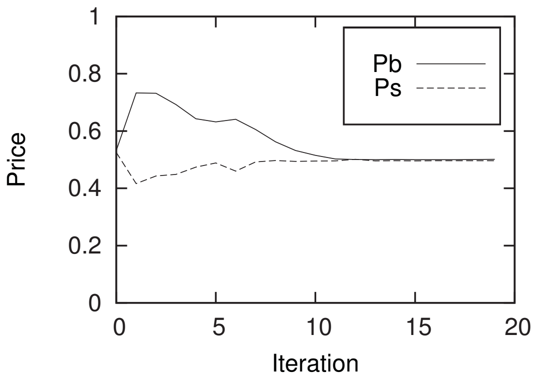
The random search algorithm was also tested when the populations of buyers and sellers were small. Specifically, $n = 1, 3, 5$ and $n = 10$ were considered as part of the experiment simulation. The plot in Figure 5–8 shows the number of times the algorithm exits if $\epsilon = 0.005$



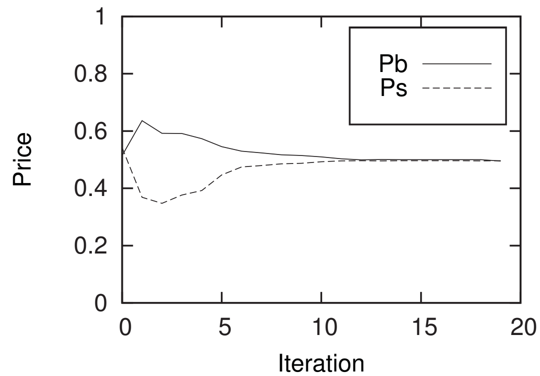
(a) GA $\mu = 0$, for utilities (5.47) and (5.48)



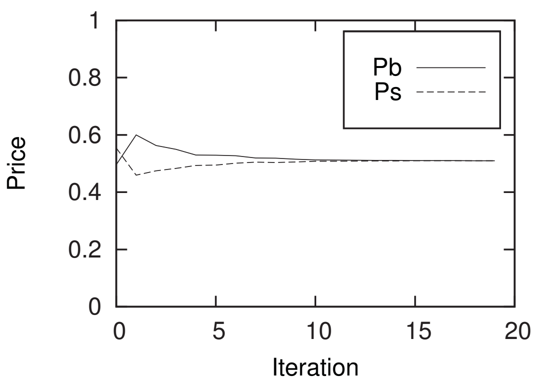
(b) GA $\mu = 0$, for utilities (5.49) and (5.50)



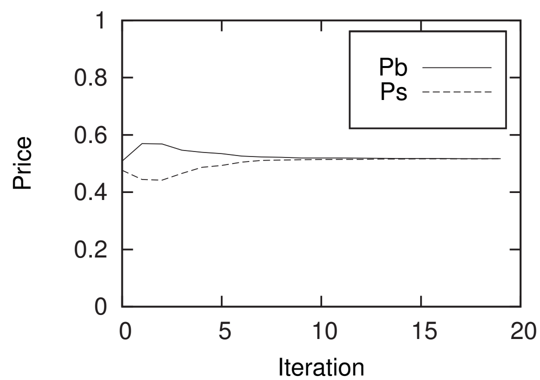
(c) GA $\mu = 0.001$, for utilities (5.47) and (5.48)



(d) GA $\mu = 0.001$, for utilities (5.49) and (5.50)

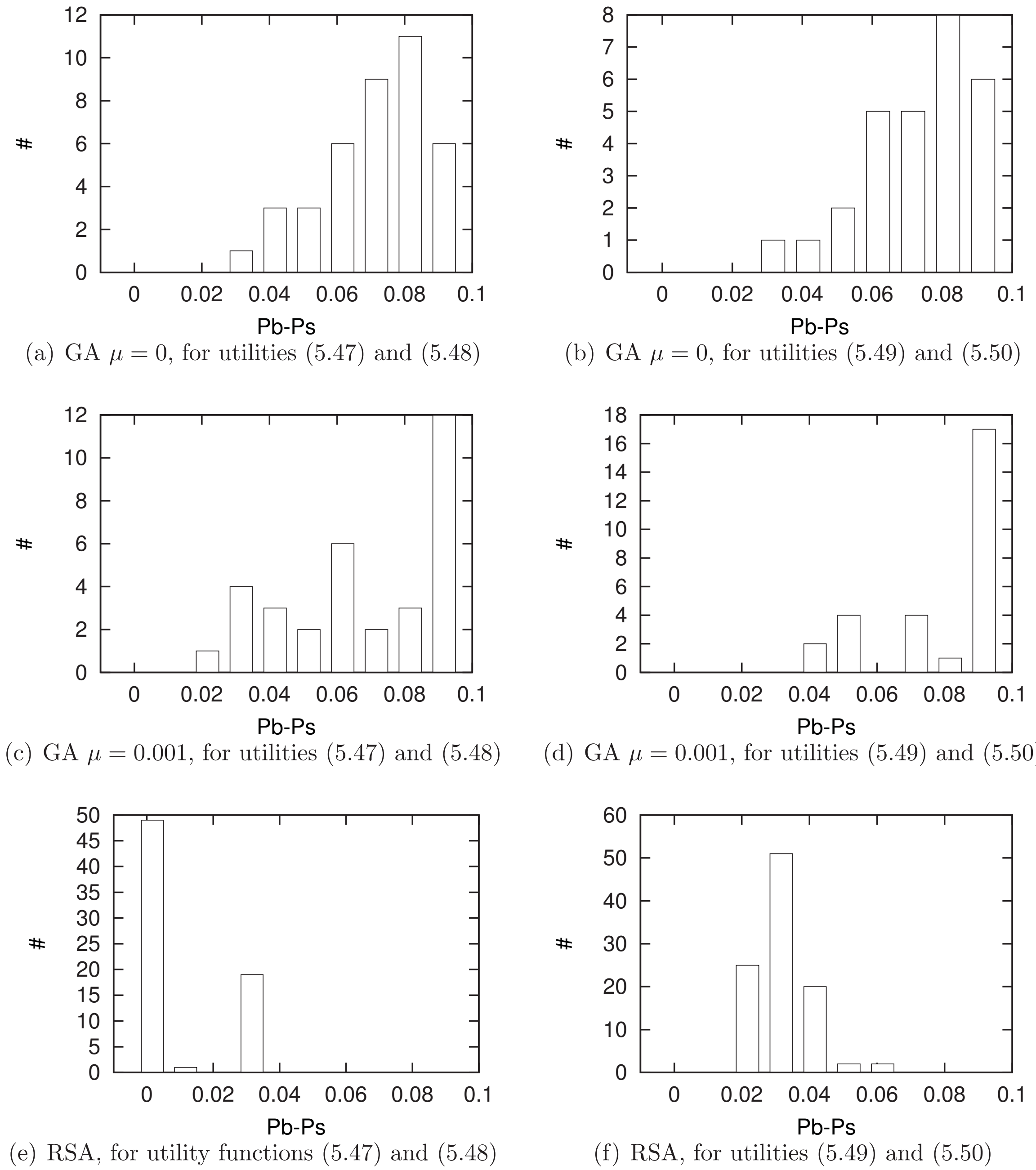


(e) RSA, for utility functions (5.47) and (5.48)



(f) RSA, for utilities (5.49) and (5.50)

Figure 5–3: Convergence of genetic and random search algorithms for 20 iterations, while using the utility functions of formulas (5.47) and (5.48) or (5.49) and (5.50). The values of parameters in Algorithm 5.1 were chosen as follows: $n = 100$, $m = 20$, $\sigma_0 = 0.1$, $\alpha = 1.2$, $\beta = 0$, and for all times $Tr = 0.5$. The values of parameters for Algorithm 5.2 are: $n = 100$, $m = 20$, crossover (Step 4 of Algorithm 5.2) for the upper half of the population, and probability of mutation as reported under sub-figures.



(a) GA $\mu = 0$, for utilities (5.47) and (5.48)

(b) GA $\mu = 0$, for utilities (5.49) and (5.50)

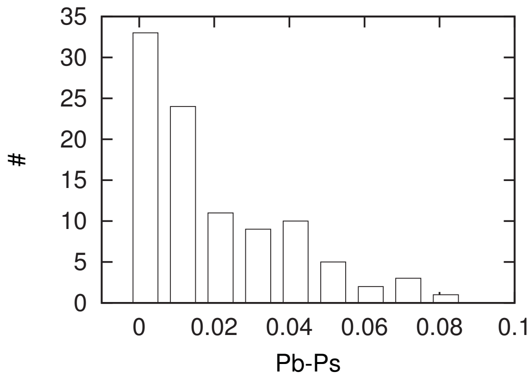
(c) GA $\mu = 0.001$, for utilities (5.47) and (5.48)

(d) GA $\mu = 0.001$, for utilities (5.49) and (5.50)

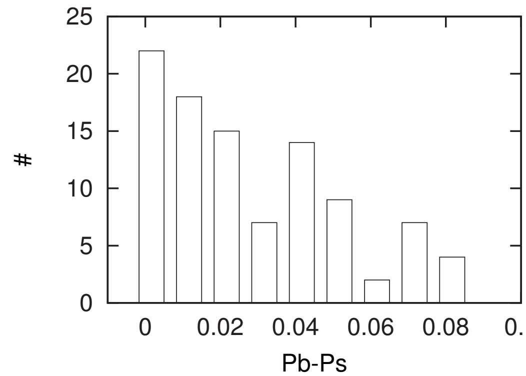
(e) RSA, for utility functions (5.47) and (5.48)

(f) RSA, for utilities (5.49) and (5.50)

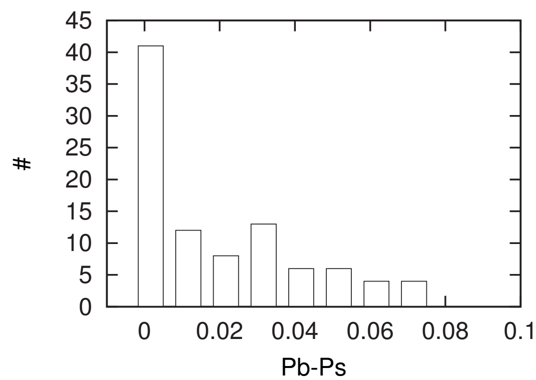
Figure 5–4: Empirical frequencies for the spread between the average bid and ask prices $p_b - p_s$ after 5 rounds, while using the utility functions of formulas (5.47) and (5.48) or (5.49) and (5.50). The values of the parameters are the same as in Figure 5–3.



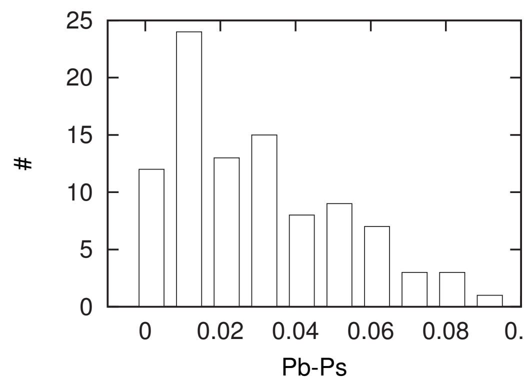
(a) GA $\mu = 0$, for utilities (5.47) and (5.48)



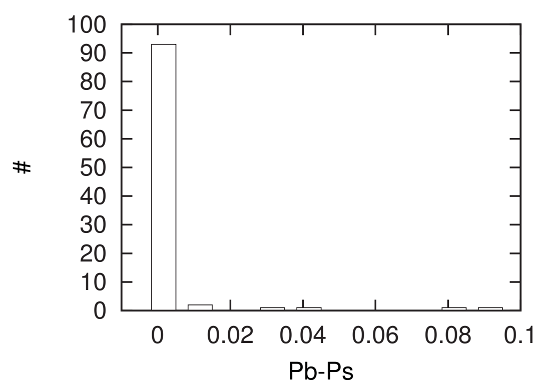
(b) GA $\mu = 0$, for utilities (5.49) and (5.50)



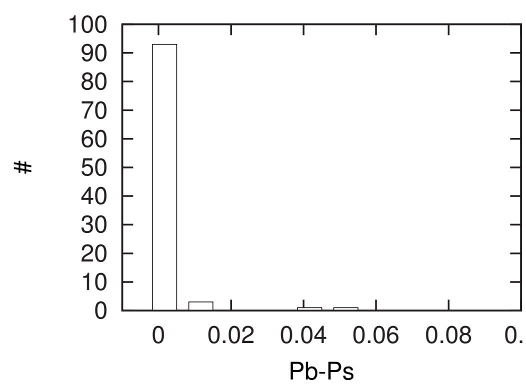
(c) GA $\mu = 0.001$, for utilities (5.47) and (5.48)



(d) GA $\mu = 0.001$, for utilities (5.49) and (5.50)

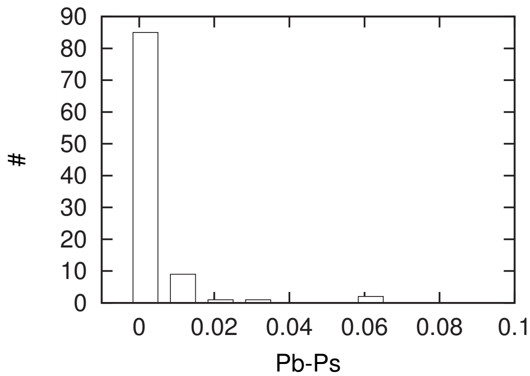


(e) RSA, for utility functions (5.47)

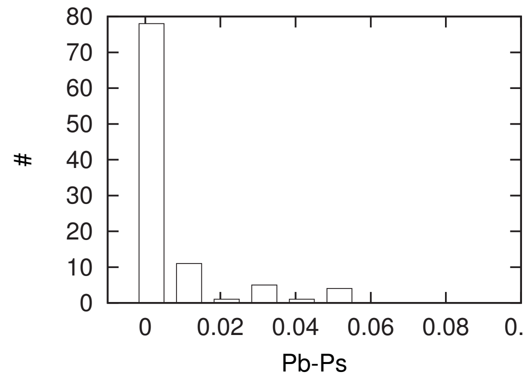


(f) RSA, for utilities (5.49 and (5.48)) and (5.50)

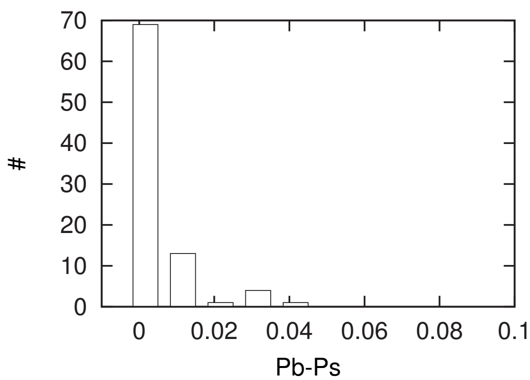
Figure 5–5: Empirical frequencies for the spread between the average bid and ask prices $p_b - p_s$ after 10 rounds, while using the utility functions of formulas (5.47) and (5.48) or (5.49) and (5.50). The values of the parameters are the same as in Figure 5–3.



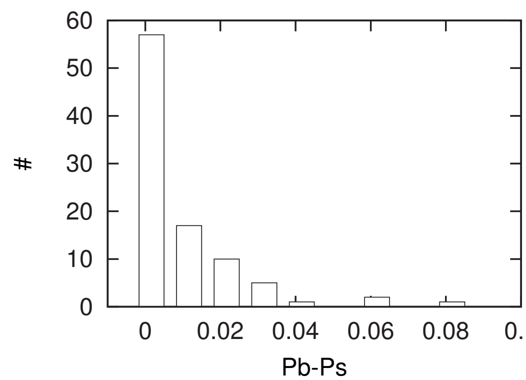
(a) GA $\mu = 0$, for utilities (5.47) and (5.48)



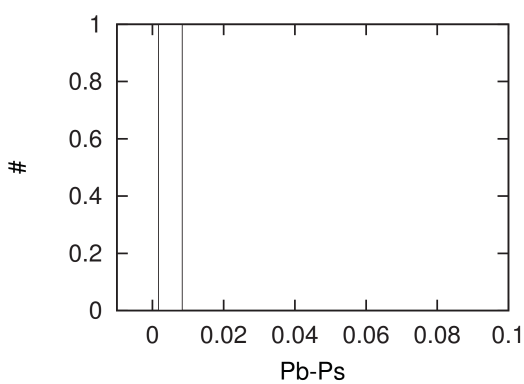
(b) GA $\mu = 0$, for utilities (5.49) and (5.50)



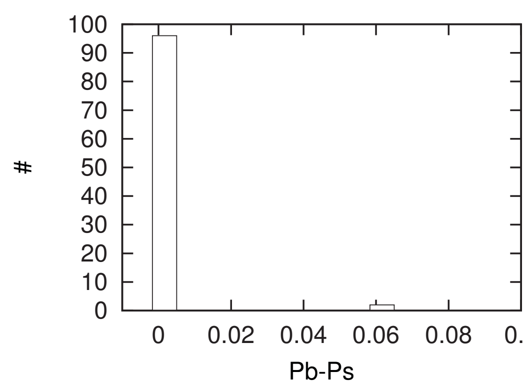
(c) GA $\mu = 0.001$, for utilities (5.47) and (5.48)



(d) GA $\mu = 0.001$, for utilities (5.49) and (5.50)

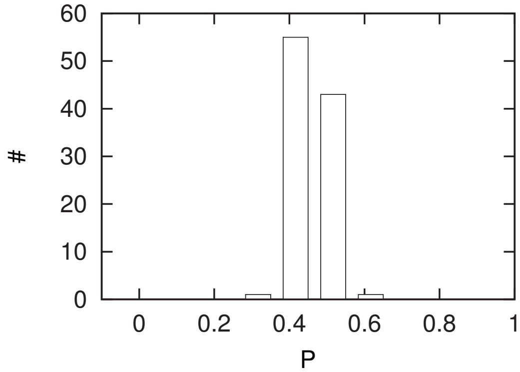


(e) RSA, for utility functions (5.47) and (5.48)

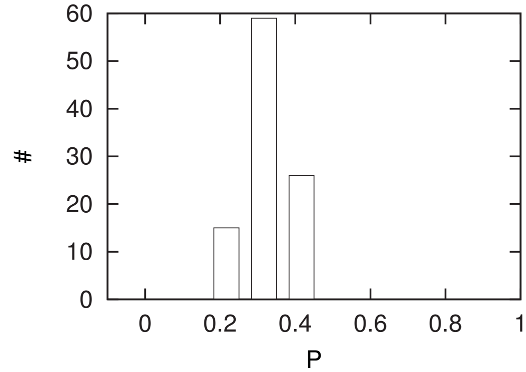


(f) RSA, for utilities (5.49) and (5.50)

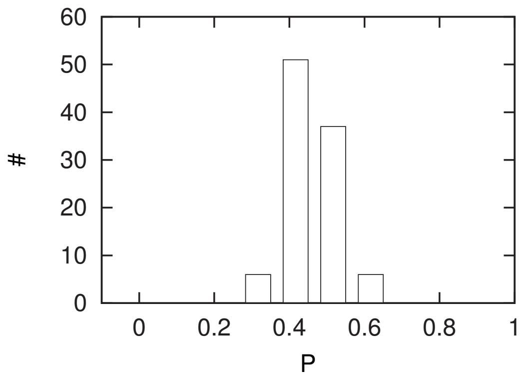
Figure 5–6: Empirical frequencies for the spread between the average bid and ask prices $p_b - p_s$ after 15 rounds, while using the utility functions of formulas (5.47) and (5.48) or (5.49) and (5.50). The values of the parameters are the same as in Figure 5–3.



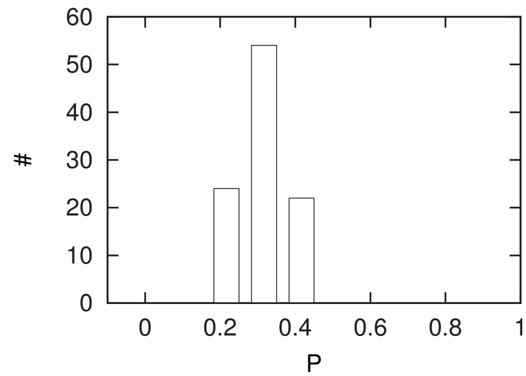
(a) GA $\mu = 0$, for utilities (5.47) and (5.48)



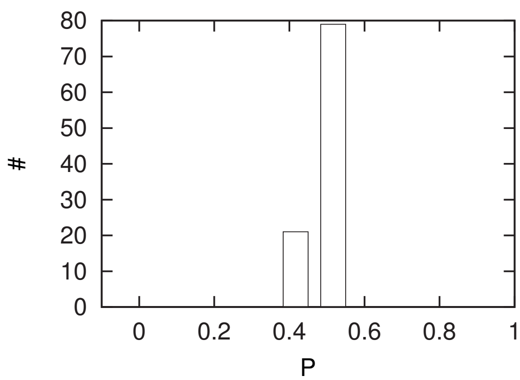
(b) GA $\mu = 0$, for utilities (5.49) and (5.50)



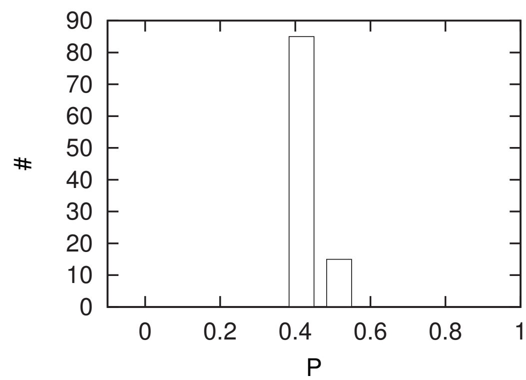
(c) GA $\mu = 0.001$, for utilities (5.47) and (5.48)



(d) GA $\mu = 0.001$, for utilities (5.49) and (5.50)



(e) RSA, for utilities (5.47) and (5.48)



(f) RSA, for utilities (5.49) and (5.50)

Figure 5–7: Choice of equilibrium after 20 rounds, while using the utility functions of formulas (5.47) and (5.48) or (5.49) and (5.50). The values the of parameters are the same as in Figure 5–3.

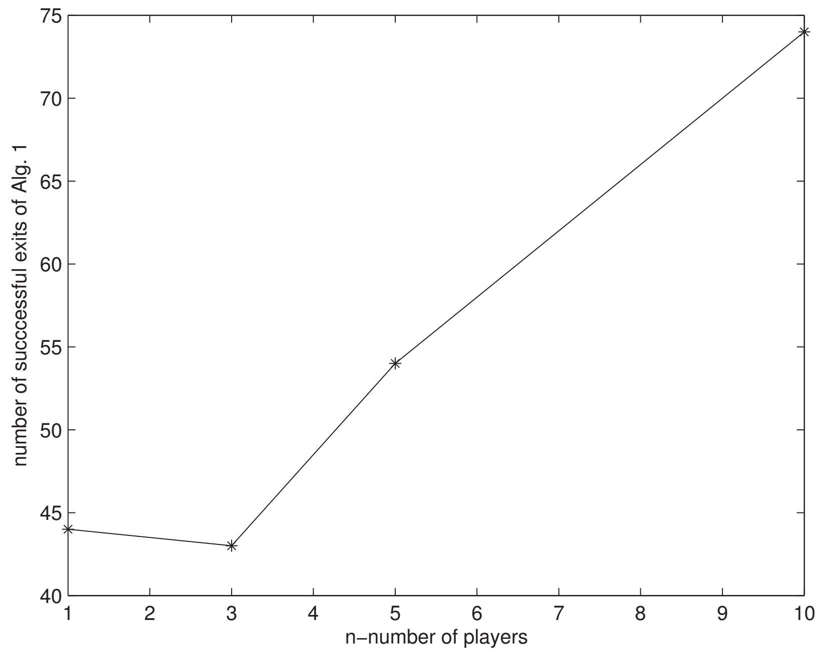


Figure 5–8: Convergence of the random search algorithm as affected by the cardinality of the populations of buyers and sellers n. The graph shows number of successful exits per 100 tests for each n. The values of the parameters are the same as in Figure 5–3.

and the algorithm was run for a 100 times. As anticipated, the exit condition is much more likely to be satisfied when the populations of players are large.

5.6 Summary

A novel stochastic algorithm which presented for learning the optimum bid in double auction markets. The mechanism of the double auction market is similar to the mechanism suggested by [91, 92], but the proposed learning algorithm tries to mimic behaviour of the sellers and buyers in a real market. In spite of the utility functions for the individual players being discontinuous, the algorithm is guaranteed to converge to a market equilibrium. Numerous simulations showed that the new algorithm converges much faster to an equilibrium than the genetic algorithm used for the same purpose [91, 92].

CHAPTER 6
Conclusion

6.1 Conclusion

The major contributions of this work are listed as follows:

In Chapter 3: Fictitious play was originally designed to calculate Nash equilibria in a bimatrix game [26]. However, Shapley showed that fictitious play may not converge to a Nash equilibrium but a polygon [30]. In the present thesis, it is first proven that in a bimatrix game, existence of a Shapley polygon is the sufficient condition for the existence of a Nash equilibrium; then this fact is used to calculate a Nash equilibrium of the game. Contrary to the time averaging method proposed by [105], the Nash equilibrium calculated by the proposed method is exact.

In Chapter 4: An evolutionary algorithm is proposed for learning in double auction markets where the buyers and sellers follow the best member of their populations from the previous round of the game and mutate their bids by a diminishing Gaussian distribution. The convergence of the stochastic learning algorithm for the risk neutral and risk averse players is proven.

In Chapter 5: A novel stochastic algorithm which presented for learning the optimum bid in double auction markets. The mechanism of the double auction market is similar to the mechanism suggested by [91, 92], but the proposed learning algorithm tries to mimic behaviour of the sellers and buyers in a real market. In spite of the utility functions for the individual players being discontinuous, the algorithm is guaranteed to converge to a market equilibrium. Numerous simulations showed that the new algorithm converges much faster to an equilibrium than the genetic algorithm used for the same purpose [91, 92].

6.2 Recommendation for Future Works

In *Chapter 3*, a method is suggested that calculates a Nash equilibrium of the bimatrix game if fictitious play has converged to a Shapley polygon. In continuation of this work, analysis of the basin of convergence for each Shapley polygon is suggested. This analysis may lead to knowledge about an upper limit on the number of Shapley polygons in a bimatrix game. Secondly, another issue which is not addressed yet is the relationship between degeneracy in a bimatrix game and the existence of Shapley polygons.

In *Chapter 4*, convergence of a certain way of learning to one of the Nash equilibria of the double auction market is studied. In continuation, one may like to know about the probabilities of selecting different equilibria, determining conditions for fluctuating behaviour in the market and convergence to the equilibria when utilities are time variant.

The algorithm presented in *Chapter 5* can be applied to other cases of auctions and discontinuous games. Further research should address the dependence of the values of the equilibria on the initial market conditions and parameters of the algorithms. A more realistic matching process should be considered in the auctioning mechanism, e.g., in real auction markets the buyers and sellers do not meet randomly, but the system selects the partners by their merits (similar to the merit based matching in the previous chapter). It is also a subject of interest to know how evolution of the bids progress, if the value of the the object in auction varies over time.

References

[1] S. Tabandeh and H. Michalska, "Random search techniques for optimal bidding in auction markets," *International Journal of Applied Mathematics*, vol. 39, pp. 276–285, 1010.

[2] S. Tabandeh and H. Michalska, "Stochastic learning of the optimum bid in auction markets," in *the 2009 International Conference of Financial Engineering*, vol. II, pp. 1325–1330, July 2009.

[3] S. Tabandeh and H. Michalska, "An evolutionary random search algorithm for double auction markets," in *Evolutionary Computation, 2009. CEC '09. IEEE Congress on*, pp. 2948–2955, May 2009.

[4] O. Morgenstern and J. Von Neumann, *Theory of Games and Economic Behavior*. Princeton University Press, May 1980.

[5] R. J. Aumann, "What is game theory trying to accomplish?," in *Frontiers of Economics* (K. Arrow and S. Honkapohja, eds.), Oxford: Basil Blackwell, 1985.

[6] I. Alger and J. W. Weibull, "Kinship, incentives, and evolution," *American Economic Review*, vol. 100, pp. 1725–58, Sept. 2010.

[7] J. W. Weibull, "Testing game theory," Working Paper Series in Economics and Finance 382, Stockholm School of Economics, May 2000.

[8] J. C. Harsanyi, "Utilities, preferences, and substantive goods," *Social Choice and Welfare*, vol. 14, no. 1, pp. 129–145, 1996.

[9] K. Arrow and F. Hahn, *General competitive analysis*. No. 6 in Mathematical economics texts, San Francisco, Calif.: Holden-Day, 1971.

[10] J. P. Benoit and V. Krishna, "Finitely repeated games," *Econometrica*, vol. 53, pp. 905–22, July 1985.

[11] M. Osborne and A. Rubinstein, *A Course in Game Theory*. Boston: The MIT Press, 1994.

[12] Y. Shoham, R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?," *Artif. Intell.*, vol. 171, no. 7, pp. 365–377, 2007.

[13] S. Sorin, "Repeated games with complete information," in *Handbook of Game Theory with Economic Applications* (R. Aumann and S. Hart, eds.), vol. 1 of *Handbook of Game Theory with Economic Applications*, ch. 4, pp. 71–107, Elsevier, 1992.

[14] D. Fudenberg and D. Levine, *The Theory of Learning in Games*, vol. 1. The MIT Press, 1 ed., 1998.

[15] J. W. Weibull, "What have we learned from evolutionary game theory so far?," Working Paper Series 487, Research Institute of Industrial Economics, September 1997.

[16] S. Demichelis and J. W. Weibull, "Language, meaning, and games: A model of communication, coordination, and evolution," *American Economic Review*, vol. 98, pp. 1292–1311, Sept. 2008.

[17] M. Benaim and M. W. Hirsch, "Mixed equilibria and dynamical systems arising from fictitious play in perturbed games," *Games and Economic Behavior*, vol. 29, pp. 36–72, Oct. 1999.

[18] J. Hofbauer, "The selection mutation equation," *Journal of Mathematical Biology*, vol. 23, pp. 41–53, Dec. 1985.

[19] W. Leinfellner, G. Eberlein, and H. Berghel, *Theory and Decision: Essays in Honor of Werner Leinfellner*. Theory and Decision Library, D. Reidel, 1988.

[20] J. Nash, "Non-cooperative games," *The Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[21] B. Stengel, "Computation of nash equilibria in finite games: introduction to the symposium," *Economic Theory*, vol. 42, pp. 1–7, Jan. 2010.

[22] R. J. Aumann, "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, vol. 1, pp. 67–96, Mar. 1974.

[23] S. Barberá, P. Hammond, and C. Seidl, *Handbook of Utility Theory: Extensions*. Handbook of Utility Theory, Kluwer Academic Publishers, 2004.

[24] G. Hanoch, "Risk aversion and consumer preferences," *Econometrica*, vol. 45, no. 2, pp. 413–26, 1977.

[25] P. R. Milgrom and R. J. Weber, "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research*, vol. 10, no. 4, pp. 619–632, 1985.

[26] G. W. Brown, "Iterative solutions of games by fictitious play," *Activity Analysis of Production and Allocation, T.C. Koopmans (Ed.), New York: Wiley*, 1951.

[27] L. Samuelson, *Evolutionary Games and Equilibrium Selection*, vol. 1 of *MIT Press Books*. The MIT Press, Oct. 1998.

[28] R. Savani and B. V. Stengel, "Hard-to-solve bimatrix games," *Econometrica*, vol. 74, no. 2, pp. 397–429, 2006.

[29] J. Robinson, "An iterative method of solving a game," *The Annals of Mathematics*, vol. 54, no. 2, pp. 296–301, 1951.

[30] L. Shapley, "Some topics in two person games," *in eds. M. Dresher et al., Advancer in Game theory.*, 1964.

[31] C. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* J. Murray, sixth ed., June 1872.

[32] R. Dawkins, *The Selfish Gene: 30th Anniversary Edition.* ISSR library, OUP Oxford, 2006.

[33] S. X. Wu and W. Banzhaf, "Investigations of wilson's and traulsen's group selection models in evolutionary computation," in *Proceedings of the 10th European conference on Advances in artificial life: Darwin meets von Neumann - Volume Part II*, ECAL'09, (Berlin, Heidelberg), pp. 1–9, Springer-Verlag, 2011.

[34] J. M. Smith, *The Theory of Evolution (Canto).* Cambridge University Press, 3 ed., July 1993.

[35] J. W. Weibull, *Evolutionary Game Theory*, vol. 1 of *MIT Press Books.* The MIT Press, Oct. 1997.

[36] U. Witt, "Evolutionary concepts in economics and biology," *Journal of Evolutionary Economics*, vol. 16, pp. 473–476, Dec. 2006.

[37] J. Hofbauer, "Evolutionary dynamics for bimatrix games: a hamiltonian system?," *Journal of Mathematical Biology*, vol. 34, no. 5-6, pp. 675–688, 1996.

[38] E. Hopkins, "Two competing models of how people learn in games," *Econometrica*, vol. 70, no. 6, pp. 2141–2166, 2002.

[39] C. Camerer and T.-H. Ho, "Experience-weighted attraction learning in normal form games," *Econometrica*, vol. 67, pp. 827–874, July 1999.

[40] E. Hopkins and R. M. Seymour, "The stability of price dispersion under seller and consumer learning," *International Economic Review*, vol. 43, pp. 1157–1190, Nov. 2002.

[41] C. M. Capra, J. K. Goeree, R. Gomez, and C. A. Holt, "Learning and noisy equilibrium behavior in an experimental study of imperfect price competition," *International Economic Review*, vol. 43, pp. 613–636, Aug. 2002.

[42] P. Milgrom and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica*, vol. 58, pp. 1255–77, Nov. 1990.

[43] N. Williams, "Stability and long run equilibrium in stochastic fictitious play. working paper," tech. rep., 2001.

[44] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning).* The MIT Press, Mar. 1998.

[45] M. Ramsza and R. M. Seymour, "Fictitious play in an evolutionary environment," *Games and Economic Behavior*, vol. 68, pp. 303–324, Jan. 2010.

[46] J. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria," *Automatic Control, IEEE Transactions on*, vol. 50, pp. 312–327, Mar. 2005.

[47] U. Berger, "Brown's original fictitious play," *Journal of Economic Theory*, vol. 135, pp. 572–578, July 2007.

[48] J. C. Harsanyi, *Games with Randomly Distributed Payoffs: A New Rationale for Mixed-strategy Equilibrium Points*. Working papers in management science, Center for Research in Management Science, University of California, Berkeley, 1972.

[49] F. Drew and D. M. Kreps, "Learning mixed equilibria," *Games and Economic Behavior*, vol. 5, pp. 320–367, July 1993.

[50] Y. Viossat, "The replicator dynamics does not lead to correlated equilibria," *Games and Economic Behavior*, vol. 59, pp. 397–407, May 2007.

[51] M. Benaïm and O. Raimond, "A class of self-interacting processes with applications to games and reinforced random walks," *SIAM J. Control Optim.*, vol. 48, pp. 4707–4730, Aug. 2010.

[52] J. J. Dai, "A note on vertex-reinforced random walks," *Statistics & Probability Letters*, vol. 62, no. 3, pp. 275–280, 2003.

[53] M. Sridharan and G. Tesauro, "Multi-agent q-learning and regression trees for automated pricing decisions," in *MultiAgent Systems, 2000. Proceedings. Fourth International Conference on*, pp. 447–448, 2000.

[54] X. Sui and H.-F. Leung, "A q-learning based adaptive bidding strategy in combinatorial auctions," in *ICEC '09: Proceedings of the 11th International Conference on Electronic Commerce*, (New York, NY, USA), pp. 186–194, ACM, 2009.

[55] C. E. Lemke and J. T. Howson, "Equilibrium points of bimatrix games," *Journal of SIAM*, vol. 12, pp. 413–423, 1964.

[56] C. H. Papadimitriou and T. Roughgarden, "Computing correlated equilibria in multi-player games," *J. ACM*, vol. 55, no. 3, pp. 1–29, 2008.

[57] N. D. Stein, P. A. Parrilo, and A. Ozdaglar, "Correlated equilibria in continuous games: Characterization and computation," *Games and Economic Behavior*, vol. 71, no. 2, pp. 436–455, 2011.

[58] A. H. van den Elzen and A. J. J. Talman, "A procedure for finding nash equilibria in bi-matrix games," *Mathematical Methods of Operations Research*, vol. 35, pp. 27–43, 1991.

[59] A. Balthasar, "Equilibrium tracing in strategic-form games," *Economic Theory*, vol. 42, pp. 39–54, Jan. 2010.

[60] S. Govindan and R. Wilson, "A global newton method to compute nash equilibria," *Journal of Economic Theory*, vol. 110, pp. 65–86, May 2003.

[61] R. Wilson, "Strategic analysis of auctions," in *Handbook of Game Theory with Economic Applications* (R. Aumann and S. Hart, eds.), vol. 1 of *Handbook of Game Theory with Economic Applications*, ch. 8, pp. 227–279, Elsevier, 1992.

[62] P. Milgrom, "Putting auction theory to work: The simultaneous ascending auction," Working Papers 98002, Stanford University, Department of Economics.

[63] R. P. McAfee and J. McMillan, "Auctions and bidding," *Journal of Economic Literature*, vol. 25, pp. 699–738, June 1987.

[64] P. R. Milgrom and R. J. Weber, "A theory of auctions and competitive bidding," *Econometrica*, vol. 50, pp. 1089–1122, Sept. 1982.

[65] M. A. Habib and A. Ziegler, "Why government bonds are sold by auction and corporate bonds by posted-price selling," fame research paper series, International Center for Financial Asset Management and Engineering, 2003.

[66] D. K. Reece, "Competitive bidding for offshore petroleum leases," *Bell Journal of Economics*, vol. 9, pp. 369–384, Autumn 1978.

[67] M. Harris and A. Raviv, "Allocation mechanisms and the design of auctions," *Econometrica*, vol. 49, pp. 1477–99, Nov. 1981.

[68] J. Niu, *Automated Auction Mechanism Design with Competing Marketplaces*. Ph.D. Dissertation, The Graduate School and University Center, The City University of New York, June 2011.

[69] G. B. Sheblé, *Computational auction mechanisms for restructured power industry operation*. Springer Netherlands, 1999.

[70] R. J. Weber, "Auction theory: By vijay krishna. academic press, 2002," *Games and Economic Behavior*, vol. 45, pp. 488–497, Nov. 2003.

[71] R. Myerson, "Optimal auction design," *Mathematics of Operations Research*, vol. 6, no. 1, pp. 58–73, 1981.

[72] N. Nisan and A. Ronen, "Algorithmic mechanism design," *Games and Economic Behavior*, vol. 35, pp. 166–196, Apr. 2001.

[73] D. Cliff, "Explorations in evolutionary design of online auction market mechanisms," *Electronic Commerce Research and Applications*, vol. 2, p. 2003, 2003.

[74] J. G. Riley and W. F. Samuelson, "Optimal auctions," *American Economic Review*, vol. 71, pp. 381–92, June 1981.

[75] S. Phelps, P. McBurney, S. Parsons, and E. Sklar, "Co-evolutionary auction mechanism design: A preliminary report," in *Revised Papers from the Workshop on Agent*

Mediated Electronic Commerce on Agent-Mediated Electronic Commerce IV, Designing Mechanisms and Systems, AAMAS '02, (London, UK), pp. 123–142, Springer-Verlag, 2002.

[76] A. Byde, "Applying evolutionary game theory to auction mechanism design," in *Proceedings of the 4th ACM conference on Electronic commerce*, EC '03, (New York, NY, USA), pp. 192–193, ACM, 2003.

[77] D. Cliff, "Zip60: further explorations in the evolutionary design of trader agents and online auction-market mechanisms," *Trans. Evol. Comp*, vol. 13, pp. 3–18, Feb. 2009.

[78] D. Cliff, "Explorations in evolutionary design of online auction market mechanisms," *Electronic Commerce Research and Applications*, vol. 2, no. 2, pp. 162–175, 2003.

[79] P. Klemperer, "Auction theory: A guide to the literature," *Journal of Economic Surveys*, vol. 13, no. 3, 1999.

[80] D. K. Gode and S. Sunder, "Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality," *Journal of Political Economy*, vol. 101, pp. 119–37, Feb. 1993.

[81] V. L. Smith, "An experimental study of competitive market behavior," *Journal of Political Economy*, vol. 70, 1962.

[82] B. Ironside, W. Joerding, and P. Kuzyk, "Double-auction market simulation software for very large classes," *Journal of Economic Education*, vol. 35, no. 3, pp. 284–289, 2004.

[83] J. I. Singer, "Double auctions across a constrained transmission line," *Oper. Res.*, vol. 50, no. 3, pp. 449–461, 2002.

[84] D. Easley and J. Ledyard, *Theories of Price Formation and Exchange in Double Oral Auctions*. Social Science working paper, California Institute of Technology. Division of the Humanities and Social Sciences, 1986.

[85] D. Fudenberg, M. Mobius, and A. Szeidl, "Existence of equilibrium in large double auctions," *Journal of Economic Theory*, vol. 133, pp. 550–567, Mar. 2007.

[86] P. Dasgupta and E. Maskin, "The existence of equilibrium in discontinuous economic games, i: Theory," *Review of Economic Studies*, vol. 53, pp. 1–26, Jan. 1986.

[87] P. Dasgupta and E. Maskin, "The existence of equilibrium in discontinuous economic games, ii: Applications," *The Review of Economic Studies*, vol. 53, no. 1, pp. 27–41, 1986.

[88] M. R. Baye, G. Tian, and J. Zhou, "Characterizations of the existence of equilibria in games with discontinuous and non-quasiconcave payoffs," *Review of Economic Studies*, vol. 60, pp. 935–48, Oct. 1993.

[89] G. Carmona, "An existence result for discontinuous games," *Journal of Economic Theory*, vol. 144, pp. 1333–1340, May 2009.

[90] D. Fudenberg, M. M. Mobius, and A. Szeidl, "Existence of equilibrium in large double auctions," Harvard Institute of Economic Research Working Papers 2033, Harvard - Institute of Economic Research, 2004.

[91] H. Dawid, "On the convergence of genetic learning in a double auction market," *Journal of Economic Dynamics and Control*, vol. 23, no. 9-10, pp. 1545–1567, 1999.

[92] H. Dawid, *Adaptive Learning by Genetic Algorithms: Analytical Results and Applications to Economic Models*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1996.

[93] S. Hon-Snir, D. Monderer, and A. Sela, "A learning approach to auctions," *Journal of Economic Theory*, vol. 82, no. 1, pp. 65–88, 1998.

[94] N. Kumarappan and P. Suriya, "Hybrid ga/sa based generation maintenance scheduling with line flow constraints in power market," in *Power Electronics, Drives and Energy Systems (PEDES) 2010 Power India, 2010 Joint International Conference on*, pp. 1–8, Dec. 2010.

[95] M. Schwind, T. Stockheim, and F. Rothlauf, "Optimization heuristics for the combinatorial auction problem," in *Evolutionary Computation, 2003. CEC '03. The 2003 Congress on*, vol. 3, pp. 1588–1595 Vol.3, Dec. 2003.

[96] S. Cincotti, E. Guerci, S. Ivaldi, and M. Raberto, "Discriminatory versus uniform electricity auctions in a duopolistic competition scenario with learning agents," in *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, 2006.

[97] K. Iyer, R. Johari, and M. Sundararajan, "Mean field equilibria of dynamic auctions with learning: A dynamic revenue equivalence theorem," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, Sept. 2011.

[98] Y. Teng, Y. Zhang, F. Niu, C. Dai, and M. Song, "Reinforcement learning based auction algorithm for dynamic spectrum access in cognitive radio networks," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, Sept. 2010.

[99] Z. Han, R. Zheng, and H. Poor, "Repeated auctions with bayesian nonparametric learning for spectrum access in cognitive radio networks," *Wireless Communications, IEEE Transactions on*, vol. 10, pp. 890–900, Mar. 2011.

[100] H. Tajima, J. Sugimoto, T. Funabashi, and R. Yokoyama, "Auction price-based thermal unit maintenance scheduling by reactive tabu search," in *Universities Power Engineering Conference, 2006. UPEC '06. Proceedings of the 41st International*, Sept. 2006.

[101] R. Gan, Q. Guo, H. Chang, and Y. Yi, "Ant colony optimization for winner determination in combinatorial auctions," in *Natural Computation, 2007. ICNC 2007. Third International Conference on*, Aug. 2007.

[102] K. Miyasawa, "On the convergence of the learning process in a 2 x 2 non-zero-sum two person game," *Economic Research Program, Princeton University, Research Memorandum No. 33.*, 1961.

[103] D. Monderer and A. Sela, "A 2 x 2 game without the fictitious play property," *Games and Economic Behavior*, vol. 14, pp. 144–148, May 1996.

[104] U. Berger, "Fictitious play in 2 x n games," *Journal of Economic Theory*, vol. 120, pp. 139–154, Feb. 2005.

[105] M. Benaïm, J. Hofbauer, and E. Hopkins, "Learning in games with unstable equilibria," *Journal of Economic Theory*, vol. 144, pp. 1694–1709, July 2009.

[106] M. Hahn, "Shapley polygons in 4 x 4 games," *Games*, vol. 1, no. 3, pp. 189–220, 2010.

[107] F. Brandt, F. Fischer, and P. Harrenstein, "On the rate of convergence of fictitious play," in *Proceedings of the Third international conference on Algorithmic game theory*, SAGT'10, (Berlin, Heidelberg), pp. 102–113, Springer-Verlag, 2010.

[108] S. Gjerstad, "The rate of convergence of continuous fictitious play," *Economic Theory*, vol. 7, no. 1, pp. 161–77, 1996.

[109] J. Hofbauer and W. H. Sandholm, "On the global convergence of stochastic fictitious play," *Econometrica*, vol. 70, pp. 2265–2294, Nov. 2002.

[110] M. Benaïm, J. Hofbauer, and S. Sorin, "Stochastic approximations and differential inclusions," *SIAM J. Control Optim.*, vol. 44, no. 1, pp. 328–348, 2005.

[111] M. Benaïm and J. Weibull, "Deterministic approximation of stochastic evolution in games," *Econometrica*, vol. 71, pp. 873–903, May 2003.

[112] A. Gaunersdorfer and J. Hofbauer, "Fictitious play, shapley polygons, and the replicator equation," *Games and Economic Behavior*, vol. 11, no. 2, pp. 279–303, 1995.

[113] D. Clarke, *Foundations of Analysis: with an Introduction to Logic and Set Theory*. Appleton-Century mathematics series, Appleton-Century-Crofts, 1971.

[114] M. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game," *Nature*, vol. 364, pp. 56–58, July 1993.

[115] M. F. Huber, *Probabilistic Framework for Sensor Management*. PhD thesis, 2009.