GAPEX: AN AGENT-BASED FRAMEWORK FOR POWER EXCHANGE MODELING AND SIMULATION

Silvano Cincotti and Giulia Gallo
DIBE-CINEF, University of Genoa, Via Opera Pia 11A, 16145 Genoa, Italy

Keywords: Agent-based computational economics, Electricity markets, Reinforcement learning, Multi-agent systems.

Abstract: The paper presents an agent-based framework for modeling and simulating power exchanges, the Genoa Artificial Power Exchange (GAPEX). The framework is implemented in MATLAB using the OOP paradigm, which allows one to define classes using a Java/C++ like syntax. GAPEX allows creation of artificial power exchanges where what-if analysis can be performed. GAPEX also reproduces exactly the market clearing procedure (e.g. by calculating Locational Marginal Prices based on the Italian high-voltage transmission network with its zonal subdivision) and the generation plants modeled are in direct correspondence with the real ones. Moreover, the presence of affine total cost functions for the generation plants results in payoff either positive, negative and null. This has major implications as negative reward are not generally considered by reinforcement learning algorithms. In order to overcome such limitation, an enhanced version of the Roth-Erev algorithm (i.e., that takes into account also negative payoffs) is presented and discussed. Results point out effectiveness of the proposed enhanced learning algorithm. Moreover, computational experiments performed within GAPEX point out a close agreement with historical real market data during both peak- and off-peak load hours thus confirming the direct applicability of GAPEX to model and to simulate power exchanges.

1 INTRODUCTION

In the last decade, large efforts have been dedicated in developing theoretical and computational approaches for modeling deregulated electricity markets. Several papers have appeared in the agent-based computational economics (hereafter ACE) literature on wholesale electricity markets and ACE has become a reference paradigm for researchers working on electricity market topics (see as reference examples (Nicolaisen et al., 2001), (Bower and Bunn, 2001), (Bunn and Oliveira, 2001), (Bagnall and Smith, 2005), (Cincotti et al., 2005), and (Sun and Tesfatsion, 2007)). Generally speaking, these papers adopt a computer-based modeling approach for studying the electricity markets as result of the interactions between heterogenous market participants. In particular, the AMES model (Agent-based Modeling of Electricity Systems, (Sun and Tesfatsion, 2007)) comprised a two-settlement system consisting of a day-ahead market and a real-time one which are both cleared by means of Locational Marginal Pricing. (Ruperez Micola et al., 2008) presented a model that consists of three sequential oligopolistic energy markets representing a wholesale gas market, a wholesale electricity market and a retail electricity market. (Weidlich and Veit, 2006) simulated two markets that are cleared sequentially, a day-ahead electricity market and a market for balancing power. (Cau and Anderson, 2002) developed a wholesale electricity market model similar to the Australian National Electricity Market. Detailed reviews on agent-based models applied to wholesale electricity markets can be found in (Weidlich and Veit, 2008b) and (Guerci et al., 2010).

In this paper, we present the Genoa Artificial Power Exchange (GAPEX), an agent-based framework for modeling and simulating electricity markets. In particular, the general GAPEX framework is presented that allows us to generalize the models and to overcome some limitations and simplifications that characterized preliminary version of the framework ((Cincotti et al., 2005), (Guerci et al., 2007) and (Rastegar et al., 2009)). In this paper, attention is devoted to model design and developing within GAPEX. This has direct implication on the features of the intelligent agents (i.e. Gencos) as well as on the mechanism of the power exchange. In particular, in order to properly model the decision process of the economic agents, an enhanced version of the classical Roth-Erev reinforcement learning algorithm (Roth and Erev, 1995) is described so to apply reinforcement learning in case of negative payoffs. Fur-

Cincotti S. and Gallo G..
GAPEX: AN AGENT-BASED FRAMEWORK FOR POWER EXCHANGE MODELING AND SIMULATION.
DOI: 10.5220/0003740300330043
In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence* (ICAART-2012), pages 33-43
ISBN: 978-989-8425-96-6

thermore, due to its complex high-voltage transmission network, the Italian power exchange (IPEX) is taken as case of study.

Results point out that GAPEX is an adequate framework to model and to simulate power exchanges. In particular, the agent-based model of the Italian Electricity Market is able to replicate market historical results during both peak- and off-peak load hours as well as to give insights on Genco behaviors. Moreover the proposed enhanced version of the classical Roth-Erev reinforcement learning algorithm points out effective learning properties with respect to existent variants in the literature.

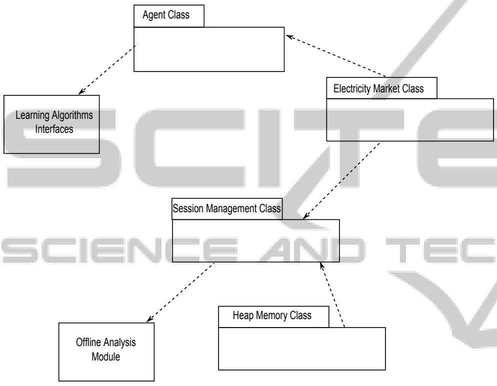

Figure 1: GAPEX Class Architecture.

The structure of the paper is as follows. In the next Section, the computational design and architecture of the GAPEX framework is presented. In Section 3 the Italian Electricity Day-Ahead Market agent-based model is described. In Section 4 the enhanced Roth-Erev reinforcement learning algorithm is presented and studied. In Section 5 we present main results of the agent-based model of the Italian Power exchange, while Section 6 summarizes main results and remarks.

2 GAPEX FRAMEWORK OVERVIEW

GAPEX is an agent-based framework developed in MATLAB that is suitable for studying the dynamic performances of many electricity markets. The simulator is implemented using OOP programming capabilities of MATLAB, which allows one to define classes using a Java/C++ like syntax thus creating a flexible and extensible ABM framework which can run local simulation and also exploit the Parallel Computing Toolbox provided with MATLAB.

Detailed computational models of the power techno-socio economic systems can be realistically simulated by means of the agent-based modeling (hereafter ABM) approach. Agents can range from entities with no cognitive function (e.g., transmission grids) to sophisticated decision makers capable of communication and learning (e.g., electricity traders). According to this research paradigm, we designed and implemented a versatile software framework for studying electricity markets. Indeed, the philosophy of the project and the modularity of its implementation provide a valuable computational framework for easily implementing other critical infrastructure systems relevant to energy markets, e.g., a natural gas market.

In order to properly address the agent behaviors in different economic environments, we have used a multi-agent learning (MAL) approach so to define appropriate algorithms able to implement sophisticated decision-making rules. This represents one of the standards in the ACE literature and some common features characterize the learning models.

The framework is composed by three main classes:

- a heap class;
- a statistical off-line analysis module;
- several algorithms and market mechanisms libraries.

Figure 1 shows the GAPEX class architecture. The Agent class is an abstract class which is extended by all agents present in GAPEX Framework. It is worth noting that the Agent class is directly extended in order to define any new types of Electricity Market Agents (e.g., Wholesalers, Energy Management Divisions, etc).

As concerning the learning algorithms, they are modeled as interfaces implemented by Gencos. Current version of GAPEX is characterized by a library of the main solutions for learning algorithms proposed in the literature (e.g., Roth-Erev algorithm, Q-Learning algorithm, Marimon-McGrattan algorithm, EWA learning and GiGa WoLF algorithm). In particular, these algorithms have been extended so to consider reward both positive, negative and null, and the features of the enhanced Roth-Erev algorithm are discussed in Section 4.

The Electricity Market class allows one to define the market clearing algorithms and it is based on the Agent class. Currently, the GAPEX allows one to simulate the Italian Day Ahead Market, the EEX spot market linking DCOPFJ Package (Sun and Tesfatsion, 2007) and the Spanish Day Ahead Market. It is worth nothing that all these algorithms are interfaces as well.

The Session class has a two-fold purposes. On the one hand, it acts as a clearing house and allows one to run several iteration of a particular simulation and to call the statistical off-line module at the end of the simulation. On the other hand, it stores all market and agent information, thus acting as a repository for all data related to energy prices and quantities both at market and at agent level (e.g. choices, propensities, etc). This feature is of crucial importance for economics application as it allows the GAPEX framework to be used as an artificial world where computational experiments can be performed. Indeed, such computational experiments are mandatory so to evaluate reproducibility of stylized facts as well as statistical properties of the self-adaptive complex system under investigation (see (Ball, 2010)). Moveover, in order to model the clearing house feature and characteristics, the mechanism of Heap memory access has been simulated and recreated into a MATLAB class. This allows one to have an online repository both for economic agents and for the electricity market agent. Thus, at the end of every simulation run, the Clearing House recall the Offline Statistical Module which carry out statistical analysis as well as visualization of the computational experiment results.

Finally, it is worth remarking that GAPEX allows direct generalization, as it is possible to create different types of agents, thus allowing the design of extremely realistic agent-based models.

3 AGENT-BASED MODELING OF THE ITALIAN ELECTRICITY DAY-AHEAD MARKET

As discussed in previous Section, GAPEX is designed as a powerful and extensible agent-based framework for electricity market modeling and simulation. Current version of GAPEX allows one to simulate different power exchange protocols, but due to its complex structure, in this paper attention is dedicated to the Italian power exchange.

It is worth remarking that a power exchange strongly differs from a stock market from both structural and behavioral point of view. From a structural point of view, the power exchange mechanism is a uniform double auction whereas the stock market one is continuous time limit order book. Furthermore, energy is not a storable good (i.e., *buy&hold* strategy are not even possible) whose consumption is contemporary to the production and is characterized by strong seasonality (i.e., daily, weekly and yearly). Moreover, from a behavioral point of view, the electricity sector is characterized by strong oligopoly (i.e., a limited and basically time-invariant number of market traders) that repeat the same game on a daily based. Theoretically speaking, such economic system seems perfect for an analytical solution based on game theory, but the dimension of the game is so high that it practically impossible to study equilibria by means of traditional game theory. Despite a first glance on analytical solutions, all these elements lead to an economic system that can be effectively studied by means of a computational approach based on learning agent, thus motivating the development of GAPEX framework for the implementation of the model of the wholesale Italian Electricity Market.

Making use of preliminary versions of GAPEX, (Cincotti et al., 2005) described and implemented an agent-based model of power exchange with a uniform price auction mechanism and a learning mechanism for the Gencos. Moreover, (Guerci et al., 2007) provided the first version of the Genoa Artificial Power Exchange and compared the discriminatory and the uniform price auction mechanism with heterogenous agents. Finally, (Rastegar et al., 2009) firstly attempted to create an agent-based model of the Italian Electricity Market, with a reduced transmission network grid and a simplified description of GenCos.

It is worth remarking that version presented and discussed in this paper of both the GAPEX and the agent-based model of the Italian electricity day-ahead market are characterized by significant extensions. Firstly, agent-based model incorporates now the exact procedure employed by the *Gestore Mercati Energetici S.p.A.* (hereafter GME) (Gestore et al., 2010b) thus overcoming the limitation of previously adopted formulation that resulted a constrained ill posed optimization procedure. Furthermore, the cognitive agents in the GAPEX mke use of the Enhanced Roth-Erev reinforcement learning algorithm (presented and discussed in Section 4), developed so to take into account payoff of any sign. There are crucial features that allowed us to calculate the energy prices based on scenarios that correctly emulate real power plants, real transmission limits and real bids.

In this Section we present the agent-based model of the Italian Electricity Day-Ahead Market (hereafter ABM IPEX Model). The Italian power exchange (IPEX) started on 1st April 2004 and is currently administrated by the Gestore Mercati Energetici S.p.A., the Italian market operator. IPEX market structure is characterized by several subsequent market sessions for both trading energy and managing critical services (e.g., reserves and real-time balancing). These are the Day-Ahead Market session - DAM, (Mercato del Giorno Prima - MGP), the Adjustment Market sessions and the Ancillary Services Market. The most important (i.e., liquid) session is the Day-Ahead Market which is organized as a non-discriminatory double-auction market where approximately 60 percent of national production is traded. The main feature of the Italian Day-Ahead Market is related to the complex high-voltage transmission network and results in a zonal splitting with both location and national energy prices.

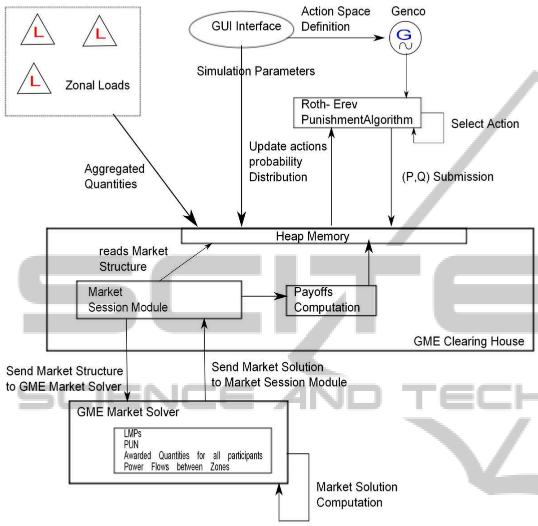

Figure 2: ABM IPEX simulation flow-diagram.

The ABM IPEX simulation flow-diagram (i.e., static representation of the objects and their interactions) is shown in Figure 2. It is worth remarking that the ABM IPEX Model consists of three main building blocks, i.e.:

- the agent-based representation of the Italian Electricity Market and the clearing mechanism regarding the Day-Ahead Market;
- the representation of the Italian Electricity Network;
- the agent-based representation of traders in the Italian Electricity Market, i.e. Gencos.

These building blocks are discussed in the following sub-sections.

3.1 Day-Ahead Market Model

GAPEX simulates Gencos bidding strategies through a daily market session in the Italian Electricity DAM. The exact market clearing procedure performed by Italian Market Operator has been implemented (see (Gestore et al., 2010a) for a detailed discussion). Furthermore, the following agents are currently represented in the model:

- Gencos: They are the economic actors at the supply side of the electricity market. They submit supply bids to the GME Market Operator and (after the market clearing procedure) they access the GME clearing house in order to retrieve market results and to update their strategic decisions. They extend GAPEX Agent class;
- Loads: They are aggregations of zonal loads and represent the demand side of the electricity market as inelastic;
- GME Market Operator: It clears the market and sends information on awarded prices and quantities to the GME clearing house. It extends GAPEX Electricity Market class;
- GME Clearing House: It computes all payoffs for the Gencos, updates their market accounts and stores all market information. It extends GAPEX Session class.

It is worth remarking that the aim of the proposed model is to represent and to study the strategic behavior of Gencos in the power exchange. Accordingly, the Gencos are characterized by sophisticated decision process (i.e., the Enhanced Roth-Erev reinforcement learning algorithm presented and discussed in Section 4)) that accounts for the effect of a repeated game. Furthermore, according to the hypothesis of a competitive electricity market, the Gencos communicate directly only with the GME Market Operator and GME Clearing House so to accounts that every Genco is only aware of its own strategies and payoffs. Finally, all the other agents in the model are passive entities and they are not endowed with any cognitive capability.

Figure 3 shows the UML class diagram for the agents modeled in the ABM IPEX:

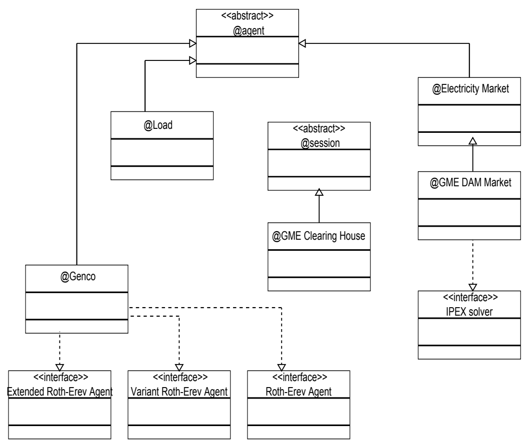

Figure 3: UML class diagram of agents in ABM IPEX.

At each iteration step, each i^{th} generator ($i = 1, 2, \ldots, N$) submits to the DAM a bidding curve shown in Figure 4. The curve is described by the triple of P_i ([€/MWh]), Q_i^- ([MWh]),Q_i^+ ([MWh]), i.e., the bidding price, the minimum and the maximum production power for i^{th} generator, respectively.

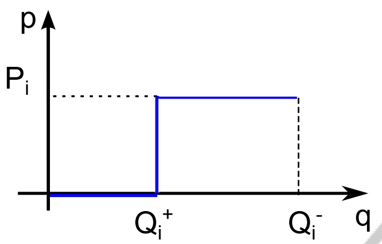

Figure 4: Reference bidding curve for a Genco.

After receiving all generators' bids, the DAM clears the market by performing a social welfare maximization subject to the constraints on the zonal energy balance (Kirchhoff's laws) and on inter-zonal transmission limits (see (Gestore et al., 2010a) for details). The objective function takes into account only the supply side of the market as the demand is assumed to be price-inelastic.

The zonal splitting clearing mechanism (i.e., DC optimal power flow procedure) allows one to determine both the unit commitment for each generator and the Locational Marginal Price (LMP) for each zone. To this aim, a graph representation of the transmission grid (that defines the area with relevant transmission constraints) is provided as input to the GAPEX (see Section 3.2). However, with respect to classical literature on power systems, the Italian market introduces two modifications. Firstly, sellers are paid at the zonal prices, i.e., Location Marginal Price (LMP), whereas buyers pay a unique national price (Prezzo Unico Nazionale - PUN) common for the whole market and computed as a weighted average of the zonal prices with respect to the zonal loads. Secondly, transmission power-flow constraints differ according to the flow direction which results in doubling the number of constraints related to the inter-zonal transmission limits. According to the specific features of the Italian market, the results of the power exchange auction consist of a set of the active powers Q_i^* and of a set of Locational Accepted Marginal Prices LMP_k for each zone $k \in \{1, 2, \ldots, K\}$.

3.2 Transmission Grid Model

The market clearing procedure described in Section 3.1 requires the definition of a transmission network. The grid structure adopted in this paper is shown in Figure 5 and reproduces the exact zonal market structure and the relative maximum transmission capacities between neighboring zones of the Italian grid model as indicated by Terna S.p.A. the Italian transmission system operator. The relevant areas of the network correspond to physical geographic areas (e.g. Northern Italy, Sicily, Sardinia, etc.) in which loads and generators, virtual production areas (i.e. foreign neighboring countries) or limited production areas (e.g. Priolo Gargallo) are present. It is worth remarking that each zone is represented as a bus to whom generators and loads are connected. Furthermore, the arches linking the zone represent the transmission connections and account for the constraints in transmissions for the power flow. Finally, transmission power-flow constraints differ according to the flow direction, e.g., power flowing from Central-South to Central-North is subject to a transmission limit that is different from the one relates to the power flowing from Central-North to Central-South. A detailed discussion of the Italian transmission grid can be found in (TERNA S.p.A., 2008).

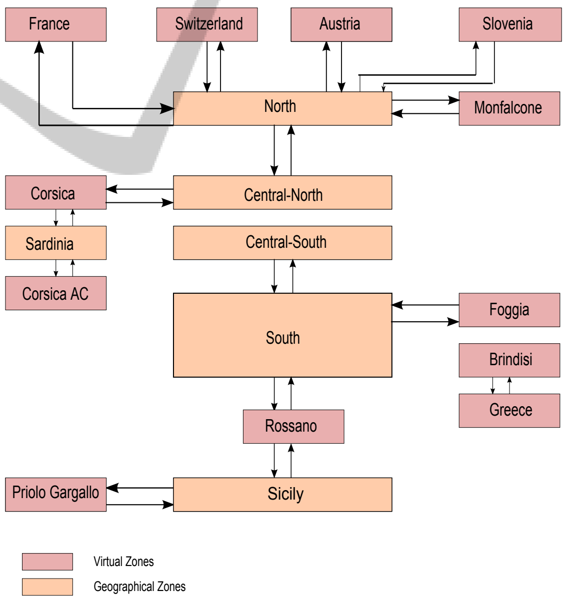

Figure 5: The Italian grid model.

3.3 Genco Model

The supply side of the market is composed by Gencos submitting bids for each of their power plants. In this paper attention is focussed on thermal power plants strategic behavior, as the remaining national production (i.e., hydro, geothermal, solar, wind) and imported production can be generally modeled as bids at zero price (Migliavacca, 2007).

A set of thermal power plants consisting of N =

175 generating units is considered. These comprise five different technologies (i.e., Coal-Fired (CF), Oil-Fired (OF), Combined Cycle (CC), Turbogas (TG) and Repower (RP)) and in the model a learning agent is associated to each generating unit.

The constant marginal costs of the i^{th} generator is assumed to be given by:

$$MC_i = \pi_i \ \ [€/\text{MWh}] \tag{1}$$

The coefficients π_i has been selected using an econometric analysis on real historical bids. The total cost function of i^{th} generator is assumed to be given by:

$$TC_i(Q_i) = a_i \cdot Q_i + b_i \ \ [€/\text{h}] \tag{2}$$

The coefficients a_i ([€/MWh]) and b_i ([€/h]) are assumed constants. a_i depends mainly on the class of efficiency and on the technology of the power plant, whereas b_i (which is specific for each power plant) accounts for investment and other quasi-fixed costs that must be recovered and that are not negligible for capital intensive industry such as the electricity one. As a consequence, the coefficients a_i have been evaluated on the basis of $MC_i(Q_i)$ with fuel costs, technology and efficiency as exogenous variables, whereas the coefficients b_i have been determined by the literature on technological business cases (Kirschen and Strbac, 2004).

Stated the cost functions of the Gencos, it is necessary to define the decision process that drives the bidding strategy. In this respect, we assume that the bidding price P_i of the i^{th} generator (see Section 3.1) is a mark-up μ_i applied to the marginal cost MC_i in equation 1, i.e.,

$$P_i = (1 + \mu_i) \cdot MC_i \tag{3}$$

As a consequence, the decision variable of the i^{th} generator is the mark-up μ_i and the learning process should individuate a profitable value for μ_i as results of the interaction (through the energy market) with the other Gencos. In particular, the profit $R_i(h)$ depends on the market clearing at hour h. Assuming that the i^{th} Genco belongs to zone k, $R_i(h)$ is given by

$$R_i(h) = LMP_k(h) \cdot Q_i^*(h) - TC_i(Q_i^*(h)) \ \ [€/\text{h}] \tag{4}$$

where TC_i is i^{th} Genco total-cost, $LMP_k(h)$ is the Location Marginal Price of zone k at hour h and $Q_i^*(h)$ is the awarded quantity to the i^{th} Genco at hour h.

Finally, it is worth remarking that the marginal cost is the reference parameter for the bids (see equation 3), whereas the total costs are crucial in order to evaluate the real profitability of the bids (see equation 4).

4 ENHANCED ROTH-EREV REINFORCEMENT LEARNING ALGORITHM

Electricity markets are characterized by inherent complexity and repeated games that requires adequate modeling of strategic behavior of traders. This is usually achieved by endowing the Gencos with learning capability. The literature on agent-based electricity market models points out three major kind of learning algorithms: *zero-intelligence* algorithms (Gode and Sunder, 1993),(Gode and Sunder, 2004), *reinforcement and belief-based* models (Camerer and Ho, 1999) and *evolutionary* approach (Nicolaisen et al., 2000).

In this paper, the strategic agent behavior is modeled by means of a reinforcement learning approach. It is worth remarking that the solutions proposed in the literature generally account for positive and null payoffs (e.g., (Nicolaisen et al., 2000) represented a first modification of the original work proposed by Roth and Erev (Roth and Erev, 1995) so to account for null payoffs). Unfortunately, this is a severe limitation in order to determine profitable strategy for economic agents in real a economic context. Indeed, the presence of fixed-costs in the cost function (see equation 2) together with market awarded quantity $Q_i^*(h) \geq 0$ for the i^{th} Genco at hour h leads to payoffs that are either positive, negative or null. This opens a question for a reinforcement learning approach that is able to cope with payoffs of any sign and to this aim we have developed an enhanced version of the Roth and Erev algorithm that is able to cope with both positive, negative and null payoffs.

The original Roth and Erev learning model (hereafter referred to as RE algorithm) considers three psychological aspects of human learning:

- the power law of practice, i.e., learning curves are initially steep and tend to progressively flatten out;
- the recency (or forgetting) effect, i.e., players recent experience plays a larger role than past experience in determining his behavior;
- the experimentation effect, i.e., not only experimented strategy but also similar strategies are reinforced.

For each strategy $a_j \in \mathcal{A}_j$ $(i = 1,..,M)$, at every round t, propensities $S_{j,t-1}(a_j)$ are updated according to:

$$S_{j,t}(a_j) = (1 - r) \cdot S_{j,t-1}(a_j) + E_{j,t}(a_j) \tag{5}$$

where $r \in [0,1]$ is the recency parameters which contributes to decrease exponentially the effect of past

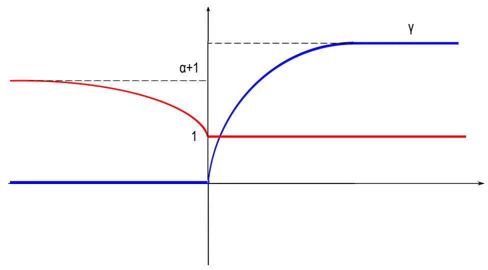

Figure 6: Functions involved in ERE. The red line shows $F[x]$, the blue one $G[x]$.

results. The second term of equation 5 is called the experimentation function and is given by:

$$E_{j,t}(a_j) = \begin{cases} \Pi_{j,t}(\hat{a}_j)\cdot(1-e) & a_j = \hat{a}_j \\ \Pi_{j,t}(\hat{a}_j)\cdot\frac{e}{M-1} & a_j \neq \hat{a}_j \end{cases} \tag{6}$$

where $e \in [0,1]$ is the experimentation parameter which assigns different weights between the played strategy and the non-played strategies and $\Pi_{j,t}(\hat{a}_j)$ is the reward obtained by playing strategy $(\hat{a}_j)$ at round t.

Propensities are then normalized so to determine the probability for the strategy selection policy $\pi_{j,t+1}(a_j)$ for the next auction round as:

$$\pi_{j,t+1}(a_j) = \frac{S_{j,t}(a_j)}{\sum_{a_j} S_{j,t}(a_j)} \tag{7}$$

The modified Roth and Erev learning model (hereafter referred to as MRE algorithm) by (Nicolaisen et al., 2000) proposed a solution for the case of zero payoffs by modifying the experimentation function in equation 6 according to:

$$E_{j,t}(a_j) = \begin{cases} \Pi_{j,t}(\hat{a}_j)\cdot(1-e) & a_j = \hat{a}_j \\ S_{j,t-1}(a_j)\cdot\frac{e}{M-1} & a_j \neq \hat{a}_j \end{cases} \tag{8}$$

It is worth remarking that MRE and RE are identical for a positive reward $\Pi_{j,t}(\hat{a}_j)$, whereas for null payoff MRE introduces an implicit *premium* for non-played strategies with respect to the ineffective (i.e. with negative $\Pi_{j,t}(\hat{a}_j)$) played strategy. MRE represents a first but not final extension of the Roth and Erev algorithm as neither MRE algorithm nor the later VRE algorithm proposed by (Sun and Tesfatsion, 2007) are able to cope with negative payoffs.

In order to overcome such limitation of the Roth-Erev algortihm, we propose to extend the MRE algorithm by enhancing the experimentation mechanism for non-played strategies according to:

$$E_{j,t}(a_j) = \begin{cases} G[\Pi_{j,t}(\hat{a}_j)]\cdot(1-e) & a_j = \hat{a}_j \\ F[\Pi_{j,t}(\hat{a}_j)]\cdot S_{j,t-1}(a_j)\cdot\frac{e}{M-1} & a_j \neq \hat{a}_j \end{cases} \tag{9}$$

where

$$G[x] = \begin{cases} -\gamma\cdot\tanh(\frac{x}{2}) & x \geq 0 \\ 0 & x \leq 0 \end{cases} \tag{10}$$

and

$$F[x] = \begin{cases} \alpha\cdot\tanh(\frac{x}{2})+1 & x \leq 0 \\ 1 & x \geq 0 \end{cases} \tag{11}$$

Figure 6 shows functions $G[...]$ and $F[...]$. It is worth noting that the proposed enhanced version represents an extension of the MRE. In particular, in the case of negative payoff, the experimentation function for the played strategies is calculated as in MRE proposed by (Nicolaisen et al., 2001) for the case of null payoffs, whereas the experimentation function of the non-played strategies is enhanced by a larger amplification the more negative is the payoff $\Pi_{j,t}(\hat{a}_j)$. This leads to an Enhanced Roth and Erev algorithm (hereafter referred to as ERE algorithm).

In the simulations discussed hereafter, we have adopted the values of 0.12 and 0.20 for the parameters e and r, respectively. Moreover, the value of 3.0 and 10.0 have been chosen for the parameters α and γ, respectively. It is worth noting that the values for e, r, α, ad γ have been chosen so to guaranty stability of the difference equations involved in the learning process (i.e., equations 5 and 9).

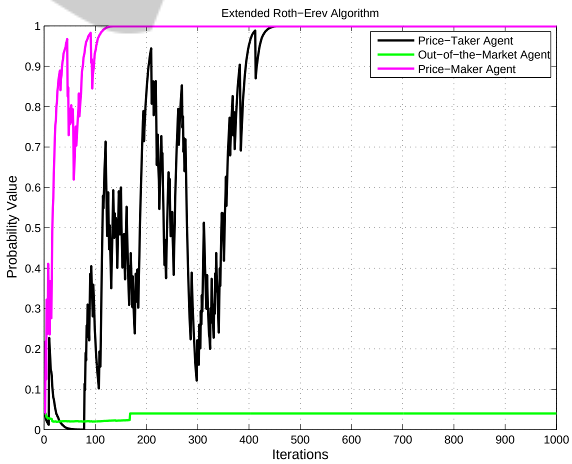

Figure 7: Convergence time-path for the different groups of interacting agents in ABM IPEX.

In order to understand effectiveness of the proposed Enhanced Roth and Erev algorithm and the interrelation between learning convergence and economic results, we firstly studied the behavior and the convergence of the learning in the power exchange model. We have assumed the initial (i.e., at $t = 0$) propensities $S_{j,t}(a_j$ in equation 5 to be uniformly distributed among the possible strategies in the strategy space. Furthermore, as discussed in Section 3.3, the strategy space is related to the mark-up variables. In all computational experiments discussed hereafter we have considered a uniformly spaced grid for μ_i in the range [0.8, 2.3] with step 0.05. This results in a set of 31 possible strategies for each of the N = 175 generators.

Stated this simulation contest, the evolution of the strategy probabilities pointed out three groups of agents:

- those whose bids are lower than clearing prices and are always accepted by the market. We denote them as price-takers agents and are characterized by a convergence of the strategy probabilities;
- those whose bids are higher than clearing prices and are always rejected by the market. We denote them as out-of-the-market agents and are generally characterized by randomly chosen strategies, as they do not participate to the market price formation and accordingly receive always negative payoffs;
- those whose bids are able to set the Locational Marginal Price. We denote them as price-maker agents and are characterized by the faster convergence time in the learning process.

Figure 7 shows an example of reference convergence time-path. For the sake of representativeness, the strategy characterized by the largest final probability (i.e., the action most willing to be played) of three reference Gencos is considered and their probabilities plotted as function of the simulation iterations.

Figure 7 points out that both price-taker and price-maker are characterized by a learning process that select the preferred action strategy (i.e., the one whose probability converge to 1). Conversely, it is worth noting that only some of the out-of-the-market agents are characterized by a convergence of the strategy probabilities. Indeed, those agents whose bids are slightly higher than the LMP tend to converge even if their bids are always rejected by the market. This can be interpreted as a result of an almost complete exploration process of their strategy spaces that allows them to conclude that the strategies played by the near competitors (i.e., the price maker agents) were characterized by a bidding price lower enough to keep them out of the market. In this exploration process, they are characterized by the slower convergence time, thus corroborating such conclusion.

5 COMPUTATIONAL EXPERIMENTS

Learning algorithms and agent-based models should stick to empirical criteria in order to demonstrate that they are able to reproduce reality. In particular, at the micro-level, learning algorithms should converge toward a price during the experiments, whereas, at the macro-level, practitioners should be able to observe stylized facts and economic emergent behaviors.

Completed the learning convergence (see Section 4), we focussed our attention to a set of computational experiments in order to understand the ability of the framework to reproduce the emergent properties shown by the IPEX DAM at macro-level.

Firstly, we have chosen a reference power exchange setting (i.e., Gencos and loads). In this respect, the scenario has been based on a real off-peak hour (i.e., hour 5 AM of Wednesday 16th December 2009) as during off-peak hour competition among producers is generally limited and thus limiting the impact on the level of prices. For the reference power exchange setting, we have performed 100 computational experiments with different random seeds in order to analyze the ensemble results of the same repeated game.

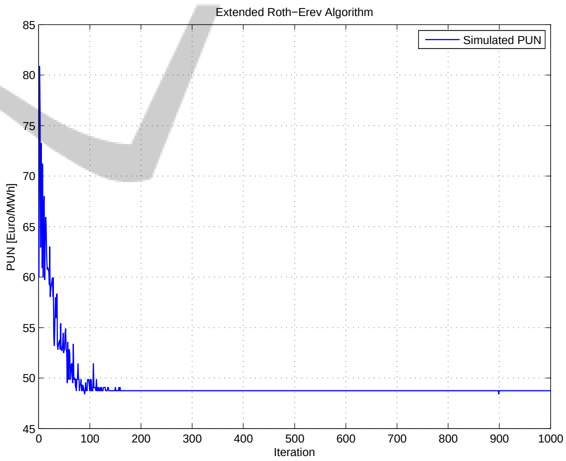

Figure 8: Convergence time-path of PUN in ABM IPEX with Enhanced Roth-Erev learning.

Both agent convergence and system convergence have been observed. While the former has been discussed in Section 4 and used as a validation proof of the enhanced Roth-Erev learning algorithm, we now concentrate on the system convergence. This type of convergence (or its lack) can be defined with respect to the convergence of the PUN time path (i.e. the clearing price converge to a value after a specific time which depends only from the participating agents). Indeed, the PUN is a weighted-average of the Locational Marginal Prices by means of the inelastic loads and an adequate representative of the market clearing and its convergence a good proxy that the system has reached an equilibrium.

Figure 8 shows a reference time-path for the PUN and it points out a convergence at the aggregate. It is worth remarking that due to the proportional update mechanism of the strategy selection probability (see Section 4), this is a real system convergence and not a fictitious one induced by a cooling parameter. We also observe that at the aggregate the learning process achieves an equilibrium that corresponds to a local optimum rather than to a global one, as the PUN and LMP dependent both on profits (i.e. payoffs) and on strategy spaces.

In order to assess the performance of the proposed Enhanced Roth-Erev algorithm, we compared it with major examples found in the literature. These are summarized in Table 1.

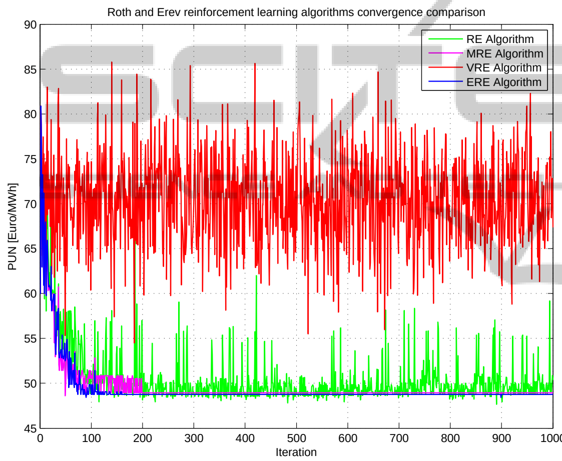

Figure 9: Convergence time-path of PUN in ABM IPEX using the learning algorithms in Table 1.

Table 1: Roth and Erev reinforcement learning algorithms.

Authors	**Formulation**
(Roth and Erev, 1995)	RE
(Nicolaisen et al., 2000) and (Weidlich and Veit, 2008a)	MRE
(Rastegar et al., 2009)	VRE
Proposed Algorithm	ERE

In this contest, we searched for the ”best performing” economic-learning algorithm, i.e. the selection of the algorithm should have lead both to learning convergence and to economic meanings.

Figure 9 shows the results. Beside the convergence at the aggregate level of the proposed ERE algorithm, results point out a convergence of the PUN only in the cases of MRE. Moreover, while ERE shows a smooth slope toward the convergence (see Figure 8), MRE points out a typical ”freezed” convergence. It is worth remarking that this is a quite ”artificial” convergence as probabilities are updated using a simulated annealing technique. Similar fictitious results have been already discussed for the Roth-Erev algorithm in a simplified agent-based electricity model (see (Jing et al., 2009)) as well as for Q-Learning (see (Watkins and Dayan, 1992)). Furthermore, in the case of VRE learning algorithm the shape of the curve suggests that although the probabilities of strategy spaces of the agents have been updated during the simulation, prices at the beginning of the simulation are the same as at the end. This directly points out that agents have not learnt any preferred strategy (i.e. there is no convergence) and leads to a ”random noise shape” of the prices, as discussed in (Jing et al., 2009).

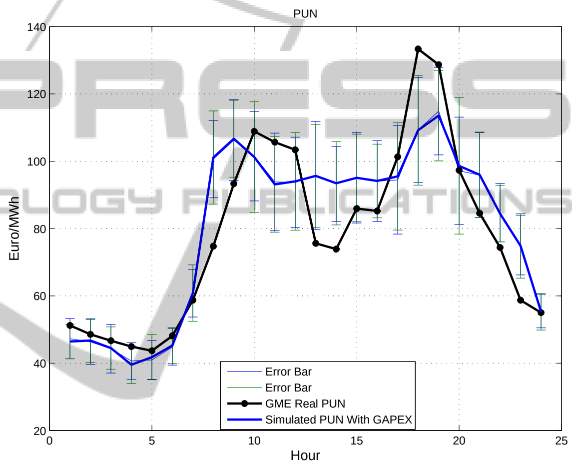

Figure 10: 24 hours GAPEX simulated PUNs vs. real GME PUNs.

It is worth remarking that these results further point out effectiveness of the proposed Enhanced Roth and Erev algorithm (with the respect to the other state-of-the-art version proposed by the literature) and its direct applicability to economic and financial context characterized by positive, null and negative rewards.

Finally, the complete 24 hours PUNs of Wednesday 16th December 2009 have been simulated. Again, we have performed 100 computational experiments (each with a length of 5,000 steps) with different random seeds in order to analyze the ensemble results of the same repeated game. It is worth remarking, that the energy market is characterized by a strong seasonality (i.e., daily, weekly, and yearly). Thus, the strategic behavior of Gencos can be properly studied on a daily base, that is the basic component of the power exchange results.

Figure 10 compares the GAPEX simulated PUNs to the GME real PUNs. Figure 10 points out that the simulated results are in good agreement throughout the whole 24 hours. Indeed, most of the GME real PUNs fall within the 95 percent (i.e., 2*σ) confidence band evaluated over the 100 computational experiment whereas the outliers are howver quite close to the limit of the 95 percent confidence band. This further states the quality and importance of the proposed methodology that is able to mostly replicate the aggregate results by means of the strategic interactions of the Gencos rather than of a black-box forecast.

It is worth remarking that these understanding the origin of the market results is a crucial element from an economics point of view as it allows us to determine the drivers and model of the power exchange. Every policy measure, antitrust action and market design requires a clear understanding of these elements in order to be effective. Furthermore, it is worth noting that in the case of the computational experiments, the generation universe is kept fixed with cost functions unchanged for the whole 24 hours. This has been assumed in order to evaluate the ability of the learning algorithm for selecting the most profitable strategy in different condition of demands. However, such condition is not present in the real GME market sessions as the generation plants are characterized by outages. The absence of outages in the computational experiments can explain the small difference between GAPEX simulated PUNs to the GME real PUNs and it is worth noting that including outages in the GAPEX is easily and direct. However, such an interesting scenario for computer science results of limited interest from an economics perspective. Indeed, it is characterized by such a large $ex-ante$ information (the exact information of the hourly participation of the Gencos to the power auction) that it results practically irrelevant and for this reason it has not been considered.

Finally, the good agreement between the GAPEX simulated PUNs and the GME real PUNs achieved by the strategic computational experiments remarks the importance of including the fixed costs in the decision-making process of Gencos. Indeed, results point out a strong relationship between fixed-costs and profits that the Enhanced Roth-Erev algorithm was able to incorporate thus improving realism of the model.

6 CONCLUSIONS

In this paper, an agent-based electricity market framework has been presented. The framework has been implemented in MATLAB using the OOP paradigm and it allows creation of artificial power exchanges characterized by real market mechanisms and by economic agent with learning capability. In order to overcome limitation in the sign of payoff typical of reinforcement learning algorithms proposed in the literature, an enhanced version of the Roth-Erev algorithm (i.e., that takes into account positive, null and negative payoffs) has been presented and discussed. Furthermore, due to its complex high-voltage transmission network, the Italian power exchange (IPEX) has been taken as case of study. This resulted in replicating the exact market clearing procedure (i.e., by calculating Locational Marginal Prices and National Price based on the Italian high-voltage transmission network with its zonal subdivision) and in considering generation plants in direct correspondence with the real ones.

Results on the convergence of the enhanced Roth-Erev learning algorithm pointed out effectiveness of the proposed solution. In particular, the evolution of the strategy probabilities pointed out different groups of agents characterized by different convergence rates that strongly depend on the role of the agent in the market. This confirms the direct applicability of the proposed Enhanced Roth-Erev learining algorithm for economic and financial applications. Moreover, computational experiments of the ABM IPEX model performed within the GAPEX pointed out a close agreement with historical data during both peak- and off-peak load hours. Thus this confirm the direct applicability of the GAPEX to model and to simulate power exchanges in particular for what-if analysis and market design.

ACKNOWLEDGEMENTS

E. Guerci and M.A. Rastegar collaborated to the design and the development of the GAPEX framework. This work has been partially supported by the University of Genoa, by the Italian Ministry of Education, University and Research (MUR) under grant PRIN 2007, by the European Social Fund (ESF) and by Regione Liguria, Italy.

REFERENCES

Bagnall, A. and Smith, G. (2005). A multi-agent model of the uk market in electricity generation. *IEEE Transactions on Evolutionary Computation 9 (5)*, pages 522–536.

Ball, P. (2010). The earth simulator. *New Scientist*, 2784:48–51.

Bower, J. and Bunn, D. W. (2001). Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the england and wales electricity market.

Journal of Economic Dynamics & Control, 25:561–592.

Bunn, D. W. and Oliveira, F. (2001). Agent-based simulation: an application to the new electricity trading arrangements of england and wales. *IEEE Transactions on Evolutionary Computation,*, 5(5):493–503. Special issue: Agent Based Computational Economics.

Camerer, C. and Ho, T. (1999). Experience-weighted attraction learning in normal-form games. *Econometrica*, 67:827–74.

Cau, T. D. H. and Anderson, E. J. (2002). A co-evolutionary approach to modelling the behaviour of participants in competitive electricity markets. *IEEE Power Engineering Society Summer Meeting*, 3:1534–1540.

Cincotti, S., Guerci, E., and Raberto, M. (2005). Agent-based simulation of power exchange with heterogeneous production companies. *Computing in Economics and Finance 2005, Society for Computational Economics*, 334.

Gestore, dei, Mercati, and Energetici (2010a). Official web site. http://www.mercatoelettrico.org/En/Default.aspx.

Gestore, del, Mercati, and Energetici (2010b). Uppo auction module user manual, appendix a - market splitting auction algorithm). Technical report, GME.

Gode, D. D. K. and Sunder, S. (2004). Double auction dynamics: structural effects of non-binding price controls. *Journal of Economic Dynamics and Control*, 28(9):1707–1731.

Gode, D. K. and Sunder, S. (1993). Allocative efficiency of markets with zero intelligence traders. market as a partial substitute for individual rationality. *J Polit Econ*, 101(1):119–137.

Guerci, E., Ivaldi, S., Raberto, M., and Cincotti, S. (2007). Learning oligopolistic competition in electricity auctions. *Computational Intelligence*, 23(2):197–220.

Guerci, E., Rastegar, M., and Cincotti, S. (2010). Agent-based modeling and simulation of competitive wholesale electricity markets. *Handbook of Power Systems*, 3(2):241–286.

Jing, Z., Ngan, H., Wang, Y., Zhang, Y., and Wang, J. (2009). Study on the convergence property of roth-erev learning model in electricity market simulation. In *Advances in Power System Control, Operation and Management (APSCOM 2009), 8th International Conference on*, pages 1 –5.

Kirschen, D. S. and Strbac, G. (2004). *Fundamentals of Power System Economics*. Wiley.

Migliavacca, G. (2007). Srems: a short-medium run electricity market simulator based on game theory and incorporating network constraints. In *Power Tech, 2007 IEEE Lausanne, Switzerland*, pages 813–818.

Nicolaisen, J., Petrov, V., and Tesfatsion, L. (2001). Market power and efficiency in a computational electricity market with discriminatory double-auction pricing. *IEEE Transactions on Evolutionary Computation*, 5(5):504–523.

Nicolaisen, J., Smith, M., Petrov, V., and Tesfatsion, L. (2000). Concentration and capacity effects on electricity market power. In *Proceedings of the 2000 Congress on Evolutionary Computation, La Jolla, USA,*, volume 2, pages 1041–1047.

Rastegar, M., Guerci, E., and Cincotti, S. (2009). Agent-based model of the Italian wholesale electricity market. In *Proceedings of the 6th International Conference on the European Energy Market EEM 09*.

Roth, A. E. and Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games Econ Behav*, 8(1):164–212.

Ruperez Micola, A., Banal Estaol, A., and Bunn, D. W. (2008). Incentives and coordination in vertically related energy markets. *Journal of Economic Behavior and Organization*, 67:381–393.

Sun, J. and Tesfatsion, L. (2007). Dynamic testing of wholesale power market designs: An open-source agent-based framework. *Comput Econ*, 30:291–327.

TERNA S.p.A. (2008). Individuazione della rete rilevante - italian version only. Technical report, TERNA S.p.A.

Watkins, C. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4):279–292.

Weidlich, A. and Veit, D. J. (2006). Bidding in interrelated day-ahead electricity markets: Insights from an agent-based simulation model. In *Proceedings of the 29th IAEE International Conference, Potsdam*.

Weidlich, A. and Veit, D. J. (2008a). Analyzing interrelated markets in the electricity sector - the case of wholesale power trading in Germany. *IEEE Power Engineering Society General Meeting*, pages 1–8.

Weidlich, A. and Veit, D. J. (2008b). A critical survey of agent-based wholesale electricity market models. *Energy Economics*, 30:1728–1759.