Evaluation of Market Rules Using a Multi-Agent System Method

Nan-Peng Yu, *Student Member, IEEE*, Chen-Ching Liu, *Fellow, IEEE*, and James Price, *Member, IEEE*

***Abstract*—The California energy crisis in 2000–2001 showed what could happen to an electricity market if it did not go through a comprehensive and rigorous testing before its implementation. Due to the complexity of the market structure, strategic interaction between the participants, and the underlying physics, it is difficult to fully evaluate the implications of potential changes to market rules. This paper presents a flexible and integrative method to assess market designs through agent-based modeling. Realistic simulation scenarios are constructed for evaluation of the proposed PJM-like market power mitigation rules of the California electricity market. Simulation results show that in the absence of market power mitigation, generation company (GENCO) agents facilitated by Q-learning are able to exploit the market flaws and make significantly higher profits relative to the competitive benchmark. The incorporation of PJM-like local market power mitigation rules is shown to be effective in suppressing the exercise of market power.**

***Index Terms*—Electricity market, market design, market power mitigation, multi-agent system, Q-learning.**

NOMENCLATURE

i	GENCO agent index.
j	LSE index.
AS_{jh}	Average per MW consumed ancillary services price charged to load serving entity j at hour h.
c_i^B	Multiplier of the supply offer for GENCO i.
c_i^{res}	Bidding price for spinning reserve capacity of unit i.
$c_i^{\mathrm{reg,up}}$	Bidding price for regulation up capacity of unit i.
$c_i^{\mathrm{reg,down}}$	Bidding price for regulation down capacity of unit i.
$C_k(h)$	LMP of real power on load bus k at hour h.
C_{jh}^G	LMP of real power at hour h for LSE j's unit.
$C_{jh}^{\mathrm{reg,up}}$	Marginal price of regulation up at hour h.
$C_{jh}^{\mathrm{reg},d}$	Marginal price of regulation down at hour h.
C_{jh}^{res}	Marginal price of spinning reserve at hour h.
$F_{\max}^l$	Thermal limit of transmission line l.
GSF_{l-k}	Generation shift factor to line l from bus k.
I	Set of GENCO agents.
L_{jh}	Total MW load of LSE j at hour h.
N_b	Number of buses in the system.
N_l	Number of lines in the system.
P_{jh}^{G*}	MW power output scheduled at hour h.
$P_{jh}^{\mathrm{reg,up}*}$	Reserved capacity for regulation up at hour h.
$P_{jh}^{\mathrm{reg},d*}$	Reserved capacity for regulation down at hour h.
$P_{jh}^{\mathrm{res}*}$	Reserved capacity for spinning reserve at hour h.
P_k	Net power injection at bus k.
P_{gk}	Total MW power generation at bus k.
P_{ih}^G	Unit i MW power generation at hour h.
P_{dk}	Total MW demand at bus k.
$P_{ih}^{\mathrm{reg,down}}$	Unit i regulation down capacity reserved at hour h.
$P_{ih}^{\mathrm{reg,up}}$	Unit i regulation up capacity reserved at hour h.
P_{ih}^{res}	Unit i spinning reserve capacity reserved at hour h.
$P_{\mathrm{Lk}}(h)$	MW load of load bus k at hour h.
R_j	Retail rates of LSE j's serving area.
R_i^{reg}	Regulation ramp rates of unit i.
R_i^{res}	Operating reserve ramp rates of unit i.
R_i^{oper}	Operational ramp rates of unit i.
$Rg_h^{\mathrm{req},d}$	System's requirement for regulation down at hour h.
$Rg_h^{\mathrm{req},u}$	System's requirement for regulation up at hour h.
Rs_h^{req}	System's requirement for spinning reserve at hour h.
τ	Delivery time requirement for ancillary service.

Manuscript received October 16, 2008; revised May 17, 2009. First published October 30, 2009; current version published January 20, 2010. This work was supported by the Power Systems Engineering Research Center (PSERC) through a collaborative project involving Iowa State University, Washington State University, and Smith College/Cornell University. Paper no. TPWRS-00781-2008.

N.-P. Yu is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50010 USA (e-mail: ynp@iastate.edu).

C.-C. Liu is with the School of Electrical, Electronic, and Mechanical Engineering, National University of Ireland, Dublin, Ireland (e-mail: liu@ucd.ie).

J. Price is with the California Independent System Operator, Folsom, CA 95630 USA (e-mail: JPrice@caiso.com).

Digital Object Identifier 10.1109/TPWRS.2009.2030379

I. INTRODUCTION

COUNTRIES around the world continue to refine their electricity market structures in various ways. There are ongoing debates over new market design issues such as how to correctly design market power mitigation (MPM) rules, how to properly implement a retail electricity market, how to effectively incorporate ancillary service (AS) markets, etc. Although much experience has been gained and costly and valuable lessons have been learned, there is still a lack of a systematic platform for evaluation of the impact of a new market design from both engineering and economic points of view. This difficulty arises from the complex interactions among strategic behaviors of market players, various layers of market designs, and the complex underlying physical network. The potential of using multi-agent system (MAS) to model complex adaptive systems has been demonstrated in various fields. Therefore, it is desirable to develop a MAS and the corresponding software platform to model the complex phenomena of an electricity market [1].

Local market power has been known as an issue for electricity markets due to limited transmission capabilities, lack of economical electricity storage devices, and short-term inelasticity of demand. During certain peak hours, electricity markets can be temporarily isolated into several subregions by N-1 and transmission thermal limit constraints. Hence, generators that possess potential local market power could leverage it to make profits through either economical or physical withholding. Furthermore, generation companies can repeatedly play in similar market scenarios and learn over time to compete less aggressively [2], [3]. Pivotal generation companies might be able to elicit collusive strategies from others by punishing uncooperative bidding behaviors. To address the problem of local market power, various types of MPM rules have been proposed and implemented in the industry. However, the effectiveness of those rules against strategic bidding market players with learning capabilities has not been extensively investigated. In general, the field of strategic bidding in an electricity market will remain an open research area for years to come.

The literature on the interaction between strategic bidding and market designs can be categorized into two approaches: equilibrium analysis and agent-based simulation. In the equilibrium analysis approach, oligopoly models such as Bertrand, Cournot, and supply function equilibrium (SFE) are used to model the stylized strategic behavior of market participants. Younes and Ilic [4] modeled the oligopolistic competition in the electricity market with SFE and Bertrand models. They recognized that inelastic load and low transmission capacities may give generators incentives to strategically constrain the network and profit from the high prices in isolated submarkets. Yao *et al.* [5] examined the two-settlement electricity market taking into account congestion, demand uncertainty, and system contingencies with a Cournot model showing that it results in lower spot equilibrium prices at most buses than a single settlement. Li and Shahidehpour [6] analyzed the strategic bidding behavior and potential market power of generation companies with SFE model. Their conclusion is that setting a lower price cap is a proper measure for mitigating market power in an electricity market. Niu *et al.* [7] modeled the electric firms' bidding behaviors with an SFE model, and studied the effects of forward contracts on the ERCOT market. They found that a high volume of forward contracts decreases the incentive of major market players to raise real-time market prices. Liu *et al.* [8] studied the impact of learning behavior of generation companies on electricity-spot-market equilibrium under repeated linear supply-function bidding. The result is that under certain conditions, the overall learning behavior will reduce market-clearing prices while in some other conditions, the results are just the contrary.

Although the equilibrium analysis yielded some useful results in the oligopoly electricity market, it may oversimplify the complicated market mechanism [9]. The accumulated bidding experience from interacting with other market participants in repeated auctions may change the perception a player has of others [10]. The advantage of a learning algorithm is that it could capture the market dynamics and provide better insights into market behaviors. In the agent-based approach, variations of reactive reinforcement learning and anticipatory reinforcement learning have been used to model the behaviors of generation companies. The learning algorithm that Bunn and Oliveira designed [9] for generators shares the same essence with reactive reinforcement learning algorithm. The average reward γ-greedy reinforcement learning (RL) method was used in [11] to model the learning and bidding processes of generation companies. These generation companies are incorporated in a nonzero sum stochastic game model to assess day-ahead (DA) market power in different auction mechanisms. The average reward γ-greedy reinforcement learning method is an RL method that uses average reward in the updating process and parameter γ to balance the exploration and exploitation. The learning configuration for generation companies in [12] is a version of stochastic reactive reinforcement learning developed by Roth and Erev [13]. A test bed was built to investigate the effects of demand-bid price sensitivity and supply-offer price caps on locational marginal prices (LMPs). Yu *et al.* [14] modeled generation companies as Q-Learning agents. The results demonstrated that Q-Learning facilitates the GENCO agent exploiting the market in the absence of an MPM process.

The use of agent-based simulation to evaluate electricity market rules has been reported in several papers [15]–[18]. Compared to the state of the art, this paper makes the following contributions.

1) This paper performs agent-based simulation on a realistic 225-bus WECC system with real heat rate data and hourly time-varying load data. The test system is created as part of this project. This work demonstrates the potential to apply multi-agent system methods to evaluate market rules of a regional market of a realistic size. Specifically, the effectiveness of PJM-like local market power mitigation (LMPM) process was fully evaluated against strategic bidding GENCO agents on the test system. The simulation results provide insights into how the MPM process suppressed the implicit collusion among the pivotal generation companies.
2) Modeling co-optimization of energy and AS market and important constraints such as transmission thermal limits, unit ramp rates, reserve, and regulation requirements in the proposed multi-agent system. This represents a step closer to reality which improves the practicality of the simulation results.
3) The proposed method considers hourly time-varying load data instead of using typical day load profiles. Load patterns and power imports vary significantly with time, which could lead to very different congestion patterns and render generators with different market power. Therefore, it is important to incorporate the hourly time varying load. To enable generation companies to learn from a dynamic market environment with time varying loads, an anticipatory reinforcement learning method, i.e., Q-learning, is used to model the bidding behavior of generation companies. Em-

powered with the Q-learning method, generation companies are capable of learning from their past bidding experience in a dynamic market environment. Furthermore, Q-learning allows generation companies to try to maximize their profit over a planning horizon rather than for one day.

4) This paper reports a software implementation of the proposed multi-agent system on an agent-oriented middleware Java Agent Development Framework (JADE) which fully complies with the Foundation for Intelligent Physical Agents (FIPA) standards. JADE is a distributed middleware system with a flexible infrastructure allowing extensions that facilitate the development of complete agent-based applications.

The remainder of this paper is organized as follows. Section II presents an MAS model for the electricity DA market (DAM). In Section III, the technical method is presented, including the software implementation of the proposed MAS. Section IV provides a study case based on the 225-bus WECC system. Section V provides the conclusion and discusses the future work.

II. Problem Formulation

An electricity DAM is composed of interacting units: market operator, generation companies, and load serving entities (LSEs). Each of them has its own goal to achieve and will not only react to changes in the market condition but also try to exert some degree of influence in the market environment. An important attribute of the DAM is that it exhibits properties arising from the interaction in the market that are not properties of the individual units themselves. Therefore, to evaluate the effectiveness of market rules of the DAM, a MAS is proposed that models the complex market dynamics among the traders. The problem formulation is motivated by CAISO's market design.

A. Multi-Agent System Structure

The DAM is modeled as a MAS with three types of interacting agents: GENCO agents, LSEs, and a market operator (MO). The DAM works as follows. Before day d begins, MO gathers the load prediction data from LSEs, and publishes the forecasted zonal load data for day $d+1$. On the morning of day d, LSEs submit their demand bids and possibly supply offers; GENCO agents submit their supply offers for DAM to MO. The MO then performs MPM and runs the market clearing software. Refer to Section II-D where details of MPM and the market clearing software are discussed. The market clearing software determines the hourly dispatch schedules to minimize the cost of purchasing energy and 100% of the AS requirement and the corresponding LMPs for energy and AS. In this MAS, MO could also perform the AS evaluation based on the market clearing results by simulating the AGC performance of the interconnected power system [19]. At the end of the process, MO sends the dispatch schedules, LMPs, and settlement information to GENCO agents and LSEs for day $d+1$.

B. GENCO Agent Model

GENCO agents sell bulk power to DAM. For simplicity, it is assumed that each GENCO agent has only one generation plant. However, this model can be extended to permit GENCO agents with multiple generation plants. Suppose the MW power output of generator i at some hour h is P_{ih}^G. For generator i, the variable production cost at hour h is represented by a quadratic form:

$$C_i\left(P_{ih}^G\right) = a_i \cdot P_{ih}^G + b_i \cdot \left(P_{ih}^G\right)^2 \tag{1}$$

where a_i, b_i are given constants. By taking derivatives on both sides of (1), the marginal cost function for generator i is obtained, i.e.,

$$\mathrm{MC}_i\left(P_{ih}^G\right) = a_i + 2 \cdot b_i \cdot P_{ih}^G. \tag{2}$$

On each day d, the GENCO agent submits to DAM a supply offer for day $d+1$ that includes two components. The first component is its reported marginal cost function given by

$$\mathrm{MC}_i\left(P_{ih}^G\right) = c_i^B\left(a_i + 2 \cdot b_i \cdot P_{ih}^G\right). \tag{3}$$

Notice that there are other alternatives to exert market power through submitting reported marginal cost functions, e.g., adding a constant term or allowing both the slope and intercept of the reported marginal cost function to be decision variables.

In this paper, it is assumed that the GENCO could exercise market power only through economical withholding. However, the modeling methodology can be extended to allow the GENCO to consider a combination of both economical and physical withholding.

The second component is its reported bidding price for AS including its bidding price for spinning reserve capacity c_i^{res}, regulation up capacity $c_i^{\mathrm{reg,up}}$, and regulation down capacity $c_i^{\mathrm{reg,down}}$. To provide regulation up or spinning reserve ancillary service, the units have to be synchronized and be able to deliver the reserved capacity within 10 min. The difference is that to provide regulation up ancillary service, the unit must be able to receive AGC signals. This is not a requirement for providing spinning reserve ancillary service. Each generator is assumed to have a set of benchmark bidding prices for AS. The reported prices of AS are calculated as the benchmark price plus a markup which was a decision variable for the generation company. There are several AS offer price markups from which GENCOs could choose. The Q-learning algorithm illustrated in Section III-B allows the GENCOs to learn from past bidding experience and to decide which markups combination is most profitable under each market condition. It is assumed that the bidding markups for spinning reserve capacity and regulation up capacity are identical for the same unit. In addition, the bidding markup for regulation down capacity is assumed to be zero. Suppose on day d, GENCO agents submit their supply offers for day $d+1$ to MO, and the market clearing program calculates LMPs for real power and AS, and dispatch schedules. Then GENCO agent i's profit on day $d+1$ is obtained by summing over the 24 h of profits on that day.

C. Load Serving Entity Model

LSEs purchase bulk power from the DAM to serve load. It is assumed that some LSEs also have generation units. If an LSE is a net buyer, then its motivation in bidding its generation would be to reduce the cost of energy and AS. Suppose the set of buses where LSE j serves loads is L_j. On day d, LSE j submits a fixed load profile for day $d+1$. The load profile specifies 24 h of MW power demand $P_{Lk}(h), h = 0, 1 \ldots 23$, at each of its load buses $k \in L_j$. Suppose LSE j submits its own generator j's reported bidding price for spinning reserve capacity c_j^{res}, regulation up capacity $c_j^{\text{reg,up}}$, regulation down capacity $c_j^{\text{reg,down}}$, and reported marginal cost function $\text{MC}_j(P_{jh}^G) = c_j^B(a_j + 2 \cdot b_j \cdot P_{jh}^G)$ to the DAM for day $d+1$. Then LSE j's profit on day $d+1$ is obtained by summing over the 24 h of profits on that day, i.e.,

$$\begin{aligned}\pi_{jD+1} &= \sum_{h=0}^{23} \begin{bmatrix} P_{jh}^{G*} C_{jh}^G + P_{jh}^{\text{reg,up}*} C_{jh}^{\text{reg,up}} \\ + P_{jh}^{\text{reg},d*} C_{jh}^{\text{reg},d} + P_{jh}^{\text{res}*} C_{jh}^{\text{res}} - C_j\left(P_{jh}^{G*}\right) \end{bmatrix} \\ &\quad + \sum_{h=0}^{23} L_{jh} R_j - \sum_{h=0}^{23} \sum_{k \in L_j} P_{\text{Lk}}(h) C_k(h) \\ &\quad - \sum_{h=0}^{23} L_{jh} \text{AS}_{jh}. \end{aligned} \tag{4}$$

The average per MW consumed ancillary services price charged to LSE j at hour t is denoted by AS_{jh}. It is calculated by dividing the total cost of procuring all ancillary services at hour h by the total amount of load at hour h.

D. *Market Operator Model*

Every day, upon receiving demand bids and supply offers, MO performs MPM and clears day-ahead energy and AS market simultaneously. The LMPM is intended to limit the exercise of local market power by generation owners in load pockets. The basic idea is to identify which generators are dispatched up to relieve congestion on noncompetitive paths (e.g., interfaces to load pockets). Generators that have been identified will be subject to mitigation since they have the potential to exercise local market power. If those generation units' supply offer is higher than default proxy bids, then energy offers will be reduced to the default level. Specifically, the MPM process includes three steps. In the first step, MO runs the market clearing software and clears the market with only competitive network constraints. In CAISO, Path 15, Path 26, Inter-ties, and interfaces to certain generation pockets are predefined as competitive network constraints. The first step is called competitive constraint run (CCR). In the second step, MO clears the market with all constraints enforced. This step is called all constraint run (ACR). In the third step, the CCR market clearing result is compared with that of the ACR. If a generation unit is incremented between CCR and ACR, the unit will be mitigated per the MPM process. In other words, mitigation applies to the units that are dispatched up by the ACR compared to the CCR. If generation unit's offer subject to mitigation is higher than cost based default proxy bids (modeled as marginal cost plus 10% in this study), then energy offers are reduced to the level of proxy bids. Those mitigated bids serve as inputs to the actual day-ahead market clearing. In reality, a method to calculate the default proxy bids is based on the unit's variable cost. Under this variable cost option, the default bids will be calculated based on the incremental heat rate curve (for gas fueled units) multiplied by the gas price index or incremental cost rate curve (for non-gas fueled units), plus an operations and maintenance adder [20]. This quantity multiplied by 110% will be used as the default proxy bid.

The market operator runs a market clearing software to determine the hourly dispatch schedules and LMPs of energy and AS. The market clearing software clears the bid-in supply with bid-in demand and procures 100% of AS requirement with minimum cost. The objective is to minimize the 24-h total purchasing cost, which is formulated as

$$\min \sum_{h=1}^{24} \left[\sum_{i \in I} \begin{pmatrix} c_i^B \left(a_i P_{ih}^G + b_i \left(P_{ih}^G\right)^2 \right) + c_i^{\text{res}} P_{ih}^{\text{res}} \\ + c_i^{\text{reg,up}} P_{ih}^{\text{reg,up}} + c_i^{\text{reg,down}} P_{ih}^{\text{reg,dwon}} \end{pmatrix} \right] \tag{5}$$

Subject to

$$P_k - P_{gk} + P_{dk} = 0, \quad k = 1, \ldots N_b \tag{6}$$

$$\left| \sum_{k=1}^{N_b} \text{GSF}_{l-k} \times P_k \right| \le F_{\max}^l, \quad l = 1, \ldots N_l \tag{7}$$

$$P_{ih}^G + P_{ih}^{\text{res}} + P_{ih}^{\text{reg,up}} \le P_i^{\max}, \quad i \in I, \quad \forall h \tag{8}$$

$$P_{ih}^G - P_{ih}^{\text{res,down}} \ge P_i^{\min}, \quad i \in I, \quad \forall h \tag{9}$$

$$0 \le \left(\frac{P_{ih}^{\text{reg,up}}}{R_i^{\text{reg}}} + \frac{P_{ih}^{\text{res}}}{R_i^{\text{res}}} \right) \le \tau, \quad i \in I, \quad \forall h \tag{10}$$

$$P_{ih}^{\text{reg,down}} \le R_i^{\text{reg}} \tau, \quad i \in I, \quad \forall h \tag{11}$$

$$\sum_{i=1}^{I} P_{ih}^{\text{reg,up}} \ge Rg_h^{\text{req},u}, \quad \forall h \tag{12}$$

$$\sum_{i=1}^{I} P_{ih}^{\text{reg,down}} \ge Rg_h^{\text{req},d}, \quad \forall h \tag{13}$$

$$\sum_{i=1}^{I} \left(P_{ih}^{\text{res}} + P_{ih}^{\text{reg,up}}\right) \ge Rs_h^{\text{req}} + Rg_h^{\text{req},u}, \quad \forall h \tag{14}$$

$$P_{ih}^G - P_{ih-1}^G \le R_i^{\text{oper}} 60, \quad i \in I, \quad \forall h \tag{15}$$

$$P_{ih-1}^G - P_{ih}^G \le R_i^{\text{oper}} 60, \quad i \in I, \quad \forall h. \tag{16}$$

The optimization problem of (5) is subject to real power balance constraints at each bus (6), thermal limit constraints for each line (7), upper and lower generation capacity constraints (8)–(9), and ramp rate constraints (10)–(11). There are also system wide reliability requirements constraints (12)–(14), and power schedule constraints between hours (15)–(16). In the case that a generation unit has reserved capacity for both regulation up and spinning reserve AS, it has to be able to deliver both within 10 min. That is why there is a combined constraint on both regulation up and spinning reserve AS in (10). In procuring upward AS, the MO could substitute a higher quality AS type to meet the requirement of a lower quality AS type if it is economically desirable to do so in the optimization process. Regulation up AS is considered to have a higher quality than spinning reserve AS. Therefore, there is an individual constraint on minimum amount of regulation up AS (12), and a combined constraint on minimum amount of both regulation up and spinning reserve AS (14). The optimization problem is solved by CPLEX which is capable of handling

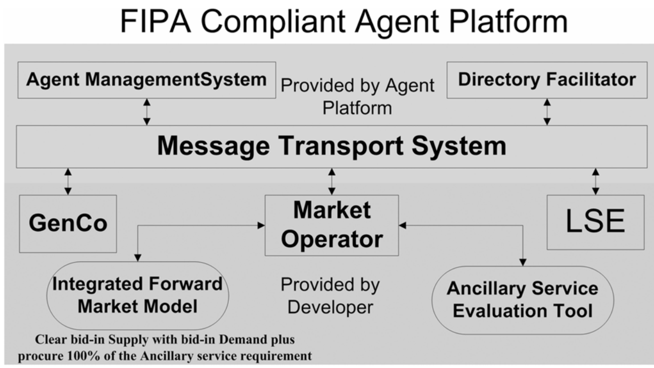

Fig. 1. Structure of the multi-agent platform for electricity DAM.

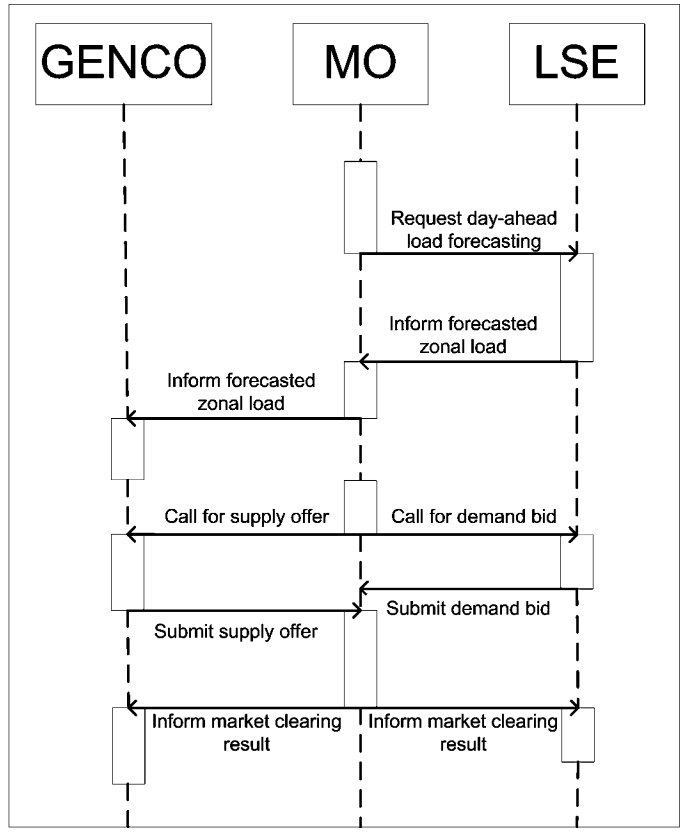

Fig. 2. Message flowing sequence in the multi-agent platform for electricity DAM.

large-scale power systems problems. A CPLEX Java interface is implemented in this project to facilitate the sharing of data between the programs.

III. Proposed Multi-Agent Approach

A. Software Implementation of Multi-Agent System

When using an agent-based approach to solve a problem, there are a number of domain independent issues that must be addressed, such as how to allow agents to communicate [21]. JADE, the most widely-used agent-oriented middleware, provides the domain independent infrastructure which allows the developers to focus on the construction of key logics. Since JADE is written in Java, it benefits from a large set of programming abstractions which greatly facilitate the development of MAS. JADE fully complies with the FIPA specifications which are maintained by the standards organization for agents and MAS. Based on the above considerations, JADE is chosen to be the middleware on which the proposed MAS was implemented.

The structure of the multi-agent platform is depicted in Fig. 1. JADE provides two utility agents: the agent management system (AMS) and directory facilitator (DF) and an inter-agent messaging system through which the agents communicate with each other. The AMS allocates agent identifiers (AIDs) to each agent that registered with it, and provides a "white page" service, where an agent can ask for the address of another. The DF provides a "yellow page" service, where agents register the services they provide, and an agent can ask for all agents to provide a particular service.

MO, GENCO agents, and LSEs are developed fully in Java in this research. Fig. 2 demonstrates the message flowing sequence in the multi-agent platform to help explain the daily sequence of tasks of MO, GENCO agents, and LSEs. A GENCO agent's daily sequence of tasks is implemented as follows: collecting forecasted zonal load data posted by MO, submitting supply offers to MO, collecting market settlement information posted by MO, and adjusting its bidding strategy based on the Q-learning algorithm. MO starts the day by collecting forecasted load data from LSEs, and posting the MO forecasted zonal load data. Upon receiving the supply offers and demand bids, it performs MPM followed by market clearing. Afterwards, it posts the market clearing information and uses an AS evaluation tool to test the system frequency performance under hypothesized disturbances. The sequence of actions taken by the LSEs is: report forecasted load data to MO, submit demand bid to MO, and collect the market settlement information from MO.

B. Learning Behavior of Agents Who Own Generation

The learning behavior of agents with generation units is modeled by Q-Learning. Q-Learning, developed by Watkins [22], is a form of anticipatory reinforcement learning that allows agents to learn how to act in a controlled Markovian domain. A controlled Markovian domain implies that the environment is Markovian in the sense that the state transition probability from any state x to another state y only depends on x, y and the action a taken by the agent, and not on the historical information. It works by successively updating estimates for the Q-values of state-action pairs. The Q-value $Q(x, a)$ is the expected discounted reward for taking action a at state x and following an optimal decision rule thereafter. The estimates of Q-values will be updated based on the reward received immediately after an action has been taken at each time step. As time moves on, series of Q-value estimates will be formed. If the series of estimates of Q-values converge to the correct Q-values, the optimal action to take in any state is the one with the highest Q-value.

The Q-learning agent moves around a discrete finite world, choosing one action from its finite action domain at every time step. In the nth step, the agent observes the current system state x_n, selects an action a_n, receives an immediate payoff r_n, and observes the next system state y_n. The agent then updates its Q-value estimates using a learning parameter α_n and a discount

factor γ [22] as follows:
If $x = x_n$ and $a = a_n$

$$Q_n(x,a) = (1-\alpha_n)Q_{n-1}(x,a) + \alpha_n[r_n + \gamma V_{n-1}(y_n)]. \tag{17}$$

Otherwise

$$Q_n(x,a) = Q_{n-1}(x,a) \tag{18}$$

where

$$V_{n-1}(y) \equiv \max_b\{Q_{n-1}(y,b)\}. \tag{19}$$

The way Q-Learning is implemented for an agent with generation unit(s) is as follows. A step in the electricity DAM environment means a trading day. The agent views the DAM as a complex system with different states. The perceived system state by an agent with generation unit(s) on day d is defined as a vector with two elements which are variables related to the zone where the agent's unit is located. The first element is predicted day $d+1$'s daily average zonal load level. The second element is the average LMP level of the most recent day that has a similar average load level as day $d+1$. Each zone's zonal daily average load is divided into $\mathrm{M_L}$ levels. For each zonal daily average load level, there are $\mathrm{M_P}$ LMP levels. Hence, the cardinality of each agent's state space is $\mathrm{M_L} \times \mathrm{M_P}$.

For an agent i, selecting an action means submitting a specific supply offer to the MO. The supply offer of the agent is defined as a vector with two elements. The first element is the bidding markup for the real power c_i^B that has $\mathrm{M_B}$ possible values. The second element is the bidding price for regulation up capacity $c_i^{\mathrm{reg,up}}$ that has $\mathrm{M_R}$ possible values. The action domain of an agent is defined as the set of all possible actions that has a dimension of $\mathrm{M_B} \times \mathrm{M_R}$. To limit the dimension of the action domain for agents, it is assumed that the bidding markups for spinning reserve capacity and regulation up capacity are identical for the same unit. In addition, the bidding markup for regulation down capacity is assumed to be zero.

The Q-learning algorithm does not specify how to choose an action at each time step. An action a in state x is selected according to the Gibbs/Boltzmann distribution given in (20) which depends on the Q-values:

$$p_D(x,a) = \frac{e^{Q(x,a)/T_d}}{\sum_{b\in \mathrm{AD}_i} e^{Q(x,b)/T_d}}. \tag{20}$$

In (20), $\mathrm{AD_i}$ is the action domain of the agent, and $\mathrm{T_d}$ is a "temperature" parameter that models a decay over time according to the formula given in Table I. In this paper, the Gibbs/Boltzmann distribution is chosen because, by setting proper parameters, it ensures a sufficient exploration while still favoring actions with higher Q-value estimates.

According to (20), when $\mathrm{T_d} = \infty$, every action has an equal probability of being chosen. As $\mathrm{T_d}$ gradually decreases over time, the action with a higher Q-value estimate will have a larger probability to be chosen. By using the Gibbs/Boltzmann distribution to select actions, the Q-learning agents are able to try a variety of actions when there was not much historical bidding information from which to learn. As time moves on, it also allows agents to progressively favor those that appear to be the best actions. In this way, the trade-off between exploration and exploitation is made.

TABLE I
Q-LEARNING PARAMETERS

γ	α	ω	$\mathrm{T_d}$	$\mathrm{M_L}$	$\mathrm{M_P}$	$\mathrm{M_B}$	$\mathrm{M_R}$
0.7	$1/T^{\omega}_{(x,a)}$	0.77	$const \times N_d^{-6}$	4	3	5	3

Consider the beginning of each day d. An agent first makes a prediction of the system state based on published load forecasting data and historical LMP data, which is represented by x. It next chooses an action according to the process illustrated above. Having chosen an action a, the agent will submit its supply offer and possibly demand bids to the MO. Once the market is cleared, the agent will receive its reward, which is the profit for day $d+1$. Then the agent uses this reward to update its Q-value estimates according to (17)–(19). In the generator model, the Q-value estimates of the state-action pairs are updated by the Q-learning algorithm.

The parameters that are used in the numerical study are set according to Table I.

In Table I, $T_{(x,a)}$ is the number of times action a has been taken in state x. N_d is the number of days that have been simulated. ω should be chosen to obtain a suitable decay for the learning parameter α. γ should be assigned a value that strikes an appropriate balance between immediate reward and expected reward in the future. The choice of these parameter values depends on the specific application. Since the application of this paper is in a dynamic multi-agent learning environment and the simulation only runs for 184 days, the γ and ω parameter are set so that the agents are able to extract enough information from the limited historical bidding experience and learn at a relatively fast pace from the environment.

IV. NUMERICAL STUDIES

A. Test System

A 225-bus WECC system developed in this project is used as the test market. The system model, which is extended from a 179-bus model used in CAISO planning studies [23], represents the essentials of the CAISO area. The system block diagram is shown in Fig. 3, where blocks with a thick dashed outline represent constrained load and generation pockets, and thick solid lines denote simplified network constraints, which are used as illustrations in CAISO's Congestion Management Reform Project, which predated market redesign and technology upgrade (MRTU).

Inside the CAISO area, 23 aggregated thermal generators are modeled as GENCO agents that bid strategically into the market. A total of 15 aggregated hydroelectric and other renewable energy generators are modeled by time-varying outputs

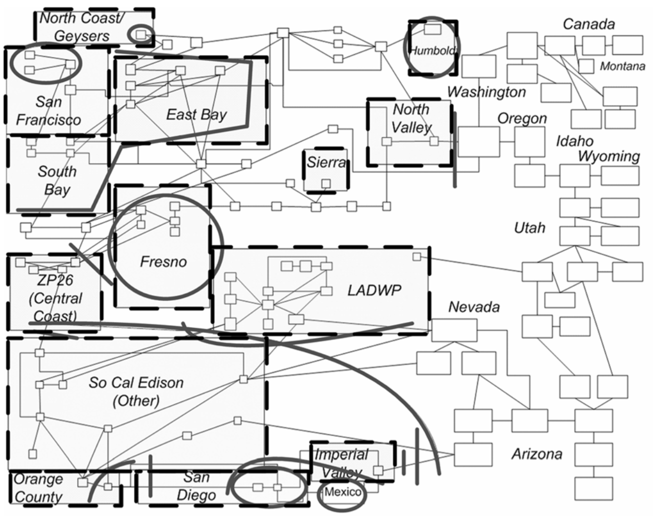

Fig. 3. The 225-bus WECC model—Details of California.

TABLE II
INSTALLED CAPACITY IN DIFFERENT AREAS OF CAISO

Area	1	2	3	4	5	6	7
Installed Capacity (MW)	4146	2644	196	1223	4010	7371	42
Area	8	9	10	11	12	13	
Installed Capacity (MW)	395	17842	3577	255	903	4669	

according to historical resource availability. Outside the CAISO area, resources represented as 22 generators produce net imports into the CAISO area. The hourly time-varying data reflect a six-month period of operations from May 1, 2004 to October 31, 2004, and include area loads for 11 local areas within the CAISO as well as net exports into a separate control area that is surrounded by the CAISO control area. The system peak demand is 44209.2 MW. The installed capacities in different areas of CAISO are listed in Table II. Due to confidentiality, names of the areas are not shown in the table.

B. Evaluation of Market Mitigation Rules of CAISO

To demonstrate the exercise of market power by Q-Learning agents and evaluate the effectiveness of the MPM rules, the following three scenarios are simulated. The first scenario is a competitive benchmark where every GENCO agent bids its marginal cost. The second scenario is an unmitigated scenario where every GENCO agent bids strategically into the market according to the Q-learning rules in the absence of MPM. The third scenario is a mitigated scenario where every GENCO agent still bids strategically into the market, but is subject to the MPM specified in Section II.

In every scenario, 15 simulation runs, each with a different random seed, are performed. The average results are reported in Figs. 4–6.

To illustrate how Q-learning facilitates the exercise of market power and implicit collusion of large GENCO agents, two pivotal GENCO agents from the SCE area are chosen for a case

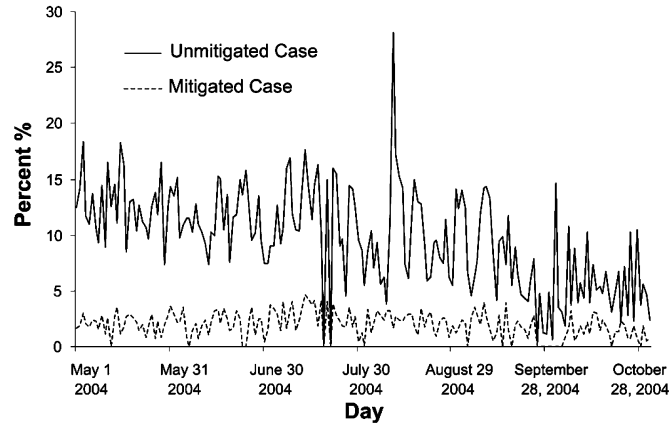

Fig. 4. Percent total market payment increase in the unmitigated and mitigated scenarios compared to the competitive benchmark.

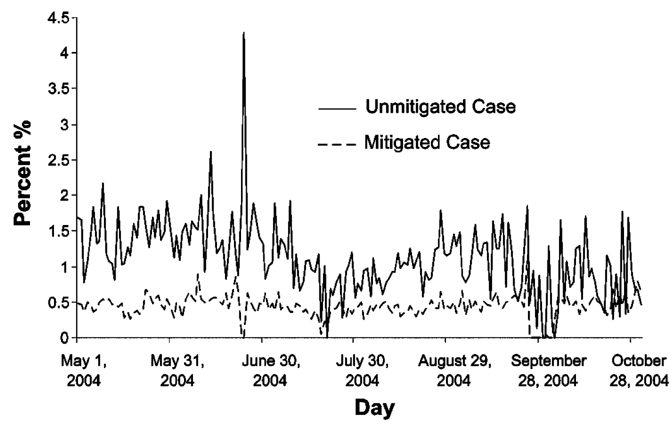

Fig. 5. Percent total generation cost increase in the unmitigated and mitigated scenarios compared to the competitive benchmark.

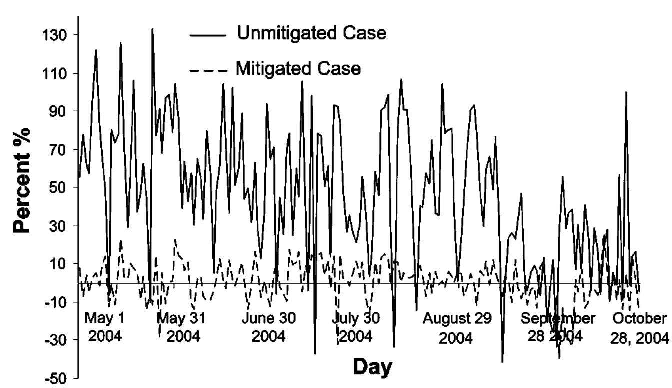

Fig. 6. Percent largest unit's profit increase in the unmitigated and mitigated scenarios compared to the competitive benchmark.

study. GENCO agents 7 and 8 together have a capacity of 7685 MW, which comprises of 64% of the area's generation capacity.

For simulation run 1 of the unmitigated scenario and mitigated scenario, key information from the Q-tables of GENCO agent 7 and 8 on August 10 are illustrated in Table III.

As can be seen from Table III, in the unmitigated scenario, both GENCO agents are in state 12. This state is encountered when the forecasted day $d+1$'s load level is high and most recent similar load level day's LMP is also high. In state 12, the highest Q-value estimate for GENCO 7 is given by action 11 which corresponds to a 12% bidding markup for real power.

TABLE III
KEY INFORMATION FROM GENCO AGENT 7 AND 8'S UPDATING Q-TABLE

	Unmitigated Scenario			Mitigated Scenario		
	State	Action Index with the Highest Estimate Q-value	Bidding Markup for Real Power	State	Action Index with the Highest Estimate Q-value	Bidding Markup for Real Power
GENCO 7	12	11	12%	12	7	8%
GENCO 8	12	7	8%	12	4	4%

Similarly, for GENCO 8, the highest Q-value estimate is given by action 7 which corresponds to an 8% bidding markup for real power. The highest possible bidding markup for real power is set to be 16% and the lowest is set to be 0%. From (20), an action that has a higher Q-value estimate will have a higher probability to be selected. Q-learning method has helped GENCO 7 and 8 to favor high markup actions when there is more potential to exercise their market power. In addition, it is shown that those two Q-learning GENCO agents are capable of implicitly colluding with each other by setting relatively high bidding markup together which will successfully drive up the price. However, the highest possible bidding markup, 16%, is not very attractive to the two pivotal GENCO agents. Indeed, although the LMPs are further driven up, they will lose part of their previously profitable generation schedule to two other relatively smaller generation companies in the area. This result extends the conclusion from [24], in that the condition of having the same demand in every trading period is not necessary. Even in a rapidly changing market environment, large generation owners who interact with each other in similar scenarios easily learn to implicitly collude even without having to know others' historical bidding data.

In the mitigated scenario, both GENCO agents are also in state 12 on Aug. 10. This time, the highest Q-value estimate for GENCO 7 is given by action 7 which correspond to an 8% bidding markup for real power. The highest Q-value estimate for GENCO 8 is given by action 4 which corresponds to a 4% bidding markup for real power. Comparing to the unmitigated case, the favorite actions' bidding markups are lower for both GENCO agents. This result shows that the MPM helped to break the high markup collusion of the two pivotal suppliers and successfully suppressed the Q-learning GENCO's potential to exert market power.

As shown in Fig. 4, the total market payment in the unmitigated scenario is significantly higher than that of the competitive benchmark. With the help of Q-Learning, the GENCO agents are able to exploit the market together and gain an average of 9.7% increase in total market payment comparing to the competitive benchmark. However, the total market payment in the mitigated scenario is slightly higher than that of the competitive benchmark. Facilitated by the MPM rules, the MO effectively reduced the percentage increase in total market payment to only 2%. The lower average load level and less congestion leads to a relatively low percentage increase of total market payment from August to the October compared to June and July.

Fig. 5 demonstrates the percentage increase of total generation cost in the mitigated and unmitigated scenario, compared to the competitive benchmark. The simulation result shows that the total generation cost increase in the unmitigated scenario is about 1.5% higher than that of the competitive benchmark. The strategic bidding of the GENCO agents' results in extra-marginal capacity being cleared, and inframarginal capacity left not dispatched. The reduction of market efficiency is caused by the market power collectively exercised by the GENCO agents. The total generation cost increase in the mitigated scenario is only about 0.5% higher than that of the competitive benchmark. This result shows that the MPM rules not only suppressed the exercise of market power but also enhanced market efficiency by bringing the total generation cost closer to marginal cost revenues, compared to the unmitigated scenario's outcome.

The largest unit's profit percentage increase in the unmitigated and mitigated scenarios, compared to the competitive benchmark, is depicted in Fig. 6. The largest GENCO agent's profit increase, which is 47.9% above the competitive benchmark, is significantly higher than the average increase of all other GENCO agents. This shows the Q-learning algorithm did help the GENCO agent realize that the huge size of its unit does provide a higher potential to exercise market power. In the mitigated scenario, the strategic bidding of generators is not beneficial to the largest GENCO agent at all. In some situations, the strategic bidding behavior will even lead to a lower profit compared to the competitive benchmark. The MPM rules being examined did reasonably well in discouraging the exercise of market power.

C. *Effects of LSE Owning Generation Resources*

It is common in agent modeling studies of electric markets to have separate agents for GENCO agents versus LSEs, and rare to have the same agents both buying and selling electricity. However, in CAISO, a number of LSEs also own or control generation. The results of this study demonstrate the importance of accounting for this type of LSE.

To examine the bidding behaviors of LSEs that own generation resources and their impacts on suppressing the GENCOs' collective market power, it is assumed that five major LSEs have their own generation units. Details of study inputs about LSEs' service areas, their units' capacity, and peak load are listed in Table IV. It is assumed that each LSE serves a peak load of twice its unit's capacity. To provide the desired test scenarios, this distribution of load among LSEs is more uniform than the actual CAISO market, in which one LSE dominates each of three transmission areas that also contain smaller municipal utilities and customers served by competitive retail energy service providers.

The simulation is carried out in four scenarios categorized by whether mitigation rules exist and whether some generation units are owned by LSEs. Fifteen simulation runs are performed in each scenario and the average results are reported below.

As shown in Fig. 7, generator 7, for example, quickly learned to bid at a lower markup in the unmitigated scenario when it is owned by an LSE and the load level is high. The LSE also learned the same strategy to reduce the cost of energy and AS in the mitigated scenario, however, at a slower rate. In the unmitigated scenario where generator 7 is owned by a GENCO agent, Q-Learning helped it learn to bid at a higher markup during high load days. In the mitigated scenario, the GENCO agent learned

TABLE IV
LSEs' DETAILED INFORMATION

	Area Peak Load (MW)	Generation Unit Owned	Unit Capacity (MW)	Peak Load to Serve (MW)
LSE A	16280.3	Generator 7	3718	7436
LSE B	16280.3	Generator 8	3967	7934
LSE C	7002.0	Generator 18	2628	5256
LSE D	6977.8	Generator 20	1478	2956
LSE E	6977.8	Generator 22	1314	2628

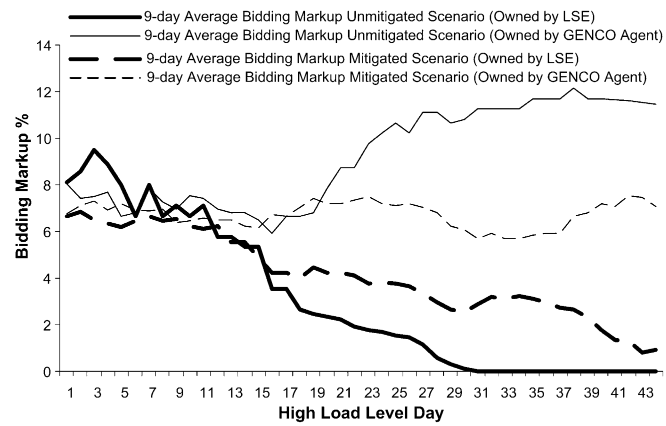

Fig. 7. Nine-day average bidding markup of generator 7 in unmitigated or mitigated scenario when owned by LSE or GENCO agent.

a similar strategy, except that the actual bidding markup cannot exceed 10% due to the existence of MPM rules. The bidding markup of other generators in Table IV also exhibits similar patterns in the four simulation scenarios.

A conclusion from these simulation results is that if the generation capacity of a LSE is smaller than the LSE's total load, it will tend to bid at its generation unit's true marginal cost. However, if a generation company owns the same generator, it will tend to bid at a much higher markup.

The total market payment and total generation cost percentage increase from competitive benchmark in the four scenarios are shown in Fig. 8. The simulation results show that both MPM procedure and the LSEs' ownership of generation units contribute to reductions in total market payment and total generation cost. In the mitigated scenario, on average, the total profit of the group of generators that are not owned by LSEs is about 1.5% lower when some LSEs own generation units compared to the case when LSEs do not own any generation units. In the unmitigated scenarios, the reduction in profit is about 1.1% on average. Therefore, the generation resources that are owned or managed by LSEs are useful for reduction of market power during peak hours to the GENCO agents.

V. CONCLUSION

This paper presents a multi-agent simulation approach to the evaluation of electricity market rules. It is found that the agent-based simulation approach empowered by Q-Learning agents is able to capture the dynamic interaction between strategic bidding market participants. The simulation result in the unmitigated scenarios shows that, even in a rapidly changing market environment, major generation owners who interact with each other in similar scenarios easily learn to implicitly collude

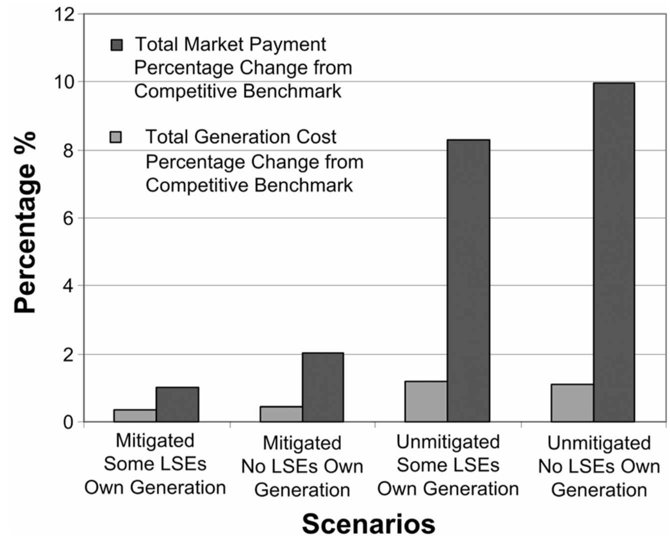

Fig. 8. Total market payment and total generation cost percentage increase in four scenarios compared to the competitive benchmark.

even without having to know others' historical bidding data. This is achieved by anticipating each other's impact on market prices. The simulation results in a mitigated scenario show that the LMPM rules proposed by CAISO perform reasonably well against Q-Learning agents and enhance the market efficiency. It is also shown that when LSEs with generation resources are net buyers in the market, they pose effectively countervailing market power against the GENCO agents. A drawback of the Q-Learning model for GENCO agents is that it may suffer from the curse of dimensionality if there are too many decision variables. This weakness can be overcome by designing a learning algorithm for electricity market participants that combines the strength of both Q-Learning and artificial neural networks.

Further research is needed on the development of the proposed multi-agent platform to enable the negotiation between GENCO agents and LSEs on bilateral contracts and study the effects of forward contracts on DAM. In addition, it is desirable to incorporate marketers into the model who trade energy but do not own generation or serve load, to examine the impacts of virtual bidding on electricity markets.

ACKNOWLEDGMENT

The authors would like to thank the project team members: Prof. A. Bose from Washington State University, Prof. J. Cardell from Smith College, and Mr. Roy from Siemens, for their helpful discussions and support. They would also like to thank Prof. G. Sheblé's for the contribution in the initial stage of this research.

REFERENCES

[1] S. D. J. McArthur *et al.*, "Multi-agent systems for power engineering applications—Part I: Concepts, approaches, and technical challenges," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1743–1752, Nov. 2007.
[2] S. Borenstein and J. Bushnell, "An empirical analysis of the potential for market power in California's electricity industry," *J. Ind. Econ.*, vol. 47, no. 3, pp. 285–323, Sep. 1999.
[3] J. B. Bushnell and F. A. Wolack, "Regulation and the leverage of local market power in the California electricity market," in *Competition Policy Center. Paper CPC00-013*, May 2, 2000. [Online]. Available: http://repositories.cdlib.org/iber/cpc/CPC00-013.

[4] Z. Younes and M. Ilic, "Generation strategies for gaming transmission constraints: Will the deregulated electric power market be an oligopoly?," *Decision Support Syst.*, vol. 24, no. 3–4, pp. 207–222, 1999.
[5] J. Yao, I. Adler, and S. Oren, "Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network," *Oper. Res.*, vol. 56, pp. 34–47, 2008.
[6] T. Li and M. Shahidehpour, "Market power analysis in electricity markets using supply function equilibrium model," *IMA J. Manage. Math.*, vol. 15, no. 4, pp. 339–354, 2004.
[7] H. Niu, R. Baldick, and G. Zhu, "Supply function equilibrium bidding strategies with fixed forward contracts," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1859–1867, Nov. 2005.
[8] Y. F. Liu, Y. X. Ni, and F. F. Wu, "Impacts of suppliers' learning behaviour on market equilibrium under repeated linear supply-function bidding," *Proc. Inst. Elect. Eng., Gen., Transm., Distrib.*, vol. 153, no. 1, pp. 44–50, Jan. 2006.
[9] D. W. Bunn and F. S. Oliveira, "Agent-based simulation: An application to the new electricity trading arrangement of England and Wales," *IEEE Trans. Evol. Comput.*, vol. 5, no. 5, pp. 493–503, Oct. 2001.
[10] M. H. Rothkopf, "Daily repetition: A neglected factor in the analysis of electricity auctions," *Elect. J.*, vol. 12, pp. 61–70, Apr. 1999.
[11] V. Nanduri and T. K. Das, "A reinforcement learning model to assess market power under auction-based energy pricing," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 85–95, Feb. 2007.
[12] H. Li, J. Sun, and L. Tesfatsion, "Effects of price-sensitive demand and price cap on LMP volatility and separation in restructured wholesale power markets," Working Paper, ISU Econ. Dept., in progress..
[13] I. Erev and A. Roth, "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria," *Amer. Econ. Rev.*, vol. 88, no. 4, pp. 848–881, 1998.
[14] N. P. Yu, C. C. Liu, and L. Tesfatsion, "Modeling of suppliers' learning behaviors in an electricity market environment," in *Proc. 14th Int. Conf. Intelligent System Applications to Power Systems (ISAP)*, Nov. 2007, pp. 1–6.
[15] I. Praça, C. Ramos, Z. Vale, and M. Cordeiro, "A new agent-based framework for the simulation of electricity markets," in *Proc. IEEE/WIC Int. Conf. Intelligent Agent Technology 2003*, Oct. 2003, pp. 469–473.
[16] D. Bunn and F. Oliveira, "Evaluating individual market power in electricity markets," *Ann. Oper. Res.*, vol. 121, no. 1–4, pp. 57–77, Jul. 2003.
[17] J. Bower and D. Bunn, "A model-based comparison of pool and bilateral market mechanisms for electricity trading," *Energy J.*, vol. 21, no. 3, pp. 1–29, Jul. 2000.
[18] M. North, G. Conzelmann, V. Koritarov, C. Macal, P. Thimmapuram, and T. Veselka, "E-laboratories: Agent-based modeling of electricity markets," in *Proc. 2002 Amer. Power Conf.*, Chicago, IL, Apr. 2002.
[19] J. Roy, A. Bose, and J. Price, "Effect of ancillary service market design on control performance of power systems," in *Proc. 2008 Power System Computation Conf. (PSCC)*.
[20] CAISO Business Practice Manual for Market Instruments, Appendix D, Calculation of Default Energy Bids, Version 1, Mar. 26, 2009. [Online]. Available: http://www.caiso.com/17ba/17baa8bc1ce20.html.
[21] F. Bellifemine, G. Caire, and D. Greewood, *Developing Multi-Agent Systems With JADE*. New York: Wiley, Feb. 2007.
[22] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, Univ. Cambridge, Cambridge, U.K., 1989.
[23] J. Jung, C. C. Liu, S. Tanimoto, and V. Vital, "Adaptation in load shedding under vulnerable operating conditions," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1199–1205, Nov. 2002.
[24] T. Mount, Investment Performance in Deregulated Markets for Electricity: A Case Study of New York State, Cornell Univ., Sep. 2007. [Online]. Available: http://www.appanet.org/files/PDFs/StudyMountEMRIreportNYISOCapacity09-07.pdf.

Nan-Peng Yu (S'06) received the B.Eng. degree in electrical engineering from Tsinghua University, Beijing, China, in 2006 and the M.S. degree in electrical engineering from Iowa State University, Ames. in 2007. He is currently pursuing the Ph.D. degree in electrical engineering at Iowa State University.

Chen-Ching Liu (F'94) received the Ph.D. degree from the University of California, Berkeley.

He was with the University of Washington, Seattle, during 1983–2005. He was Palmer Chair Professor at Iowa State University, Ames, during 2006–08. Currently, he is serving as a Professor and Deputy Principal of the College of Engineering, Mathematical, and Physical Sciences, University College Dublin, Dublin, Ireland.

Dr. Liu is the Past Chair of the IEEE PES Technical Committee on Power System Analysis, Computing, and Economics.

James Price (M'06) received the B.S. degree from the California Institute of Technology, Pasadena, and M.S. and Ph.D. degrees from Stanford University, Stanford, CA.

He is Lead Engineering Specialist in Market and Product Development at the CAISO, Folsom, CA.