Bidding in Smart Grid PDAs: Theory, Analysis and Strategy

by

Susobhan Ghosh, Sujit Prakash Gujar, Praveen Paruchuri, Easwar Subramanian, Sanjay P. Bhat

in

AAAI Conference on Artificial Intelligence

Report No: IIIT/TR/2020/-1

Centre for Visual Information Technology
International Institute of Information Technology
Hyderabad - 500 032, INDIA
January 2020
Bidding in Smart Grid PDAs: Theory, Analysis and Strategy

Conference Paper · February 2020

5 authors, including:

Susobhan Ghosh
Singapore Management University
6 PUBLICATIONS 3 CITATIONS
SEE PROFILE

Sanjay P Bhat
Tata Consultancy Services Limited
80 PUBLICATIONS 6,337 CITATIONS
SEE PROFILE

Easwar Subramanian
Tata Consultancy Services Limited
18 PUBLICATIONS 107 CITATIONS
SEE PROFILE

Sujit Gujar
International Institute of Information Technology, Hyderabad
75 PUBLICATIONS 265 CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

- General theory of dynamical systems View project
- Artificial Intelligence, Multiagent Systems, Game Theory View project
Bidding in Smart Grid PDAs: Theory, Analysis and Strategy

Susobhan Ghosh, Sujit Gujar, Praveen Paruchuri, Easwar Subramanian, Sanjay P. Bhat

1 Machine Learning Lab, IIIT Hyderabad, India {susobhan.ghosh@research.iiit.ac.in}, {sujit.gujar, praveen.p}@iiit.ac.in
2 Tata Consultancy Services, TCS Innovation Labs, Hyderabad, India {easwar.subramanian, sanjay.bhat}@tcs.com

Abstract

Periodic Double Auctions (PDAs) are commonly used in the real world for trading, e.g. in stock markets to determine stock opening prices, and energy markets to trade energy in order to balance net demand in smart grids, involving trillions of dollars in the process. A bidder, participating in such PDAs, has to plan for bids in the current auction as well as for the future auctions, which highlights the necessity of good bidding strategies. In this paper, we perform an equilibrium analysis of single unit single-shot double auctions with a certain clearing price and payment rule, which we refer to as ACPR, and find it intractable to analyze as number of participating agents increase. We further derive the best response for a bidder with complete information in a single-shot double auction with ACPR. Leveraging the theory developed for single-shot double auction and taking the PowerTAC wholesale market as our testbed, we proceed by modeling the PDA of PowerTAC as an MDP. We propose a novel bidding strategy, namely MDPLCPBS. We empirically show that MDPLCPBS follows the equilibrium strategy for double auctions that we previously analyze. In addition, we benchmark our strategy against the baseline and the state-of-the-art bidding strategies for the PowerTAC wholesale market PDAs, and show that MDPLCPBS outperforms most of them consistently.

Introduction

Auctions are mechanisms which facilitate buying and selling of goods or items amongst a group of agents. Double auctions are prevalent when both the sides of a market actively bid. For example, in the New York Stock Exchange, opening prices are determined using double auctions (Parsons, Rodriguez-Aguilar, and Klein 2011). In smart grids, multiple power generating companies and different distributing agencies (brokers) trade electricity in the wholesale markets using double auctions.

In this work, we focus primarily on electricity markets. In July 2019, approximately 1.2 Billion Euros worth electricity was traded in Nord Pool alone, with 52% of the volume being traded using APIs (Nord Pool AS 2019). Any small improvement in cost optimization by deploying better bidding strategies can lead to significant improvements in the profits of the distributing agencies. Motivated by this, we take up a formal game-theoretic approach in this work for devising bidding strategies.

Typically, for double auctions, clearing price and payment rules differ from market to market. Equilibrium analysis of double auctions has been explored extensively with different payment and clearing price rules (Wilson 1992). Specifically, for k-double auctions, Satterthwaite and Williams (1989) proved the existence of multiple non-trivial equilibria for $k \in [0,1]$. They also focused on a class of well-behaved equilibria, by making generalist assumptions on buyer’s and seller’s bidding strategies. Our focus in this paper is Average Clearing Price Rule (ACPR) based Periodic Double Auctions (PDAs), commonly used in smart grids (Power TAC (Ketter, Collins, and Weerdt 2017)). In ACPR, the clearing price set as the average of last executing bid and last executing ask (a special case of k-double auction with $k = 0.5$).

For ACPR, Chatterjee and Samuelson (1983) constructed a symmetric equilibrium for the case of one buyer and one seller with uniformly distributed valuations. However, in the vast literature of double auctions, a generic equilibrium analysis for ACPR with more buyers has not been well explored (Wilson 1992). We take up a double auction with ACPR as a case study. We assume all the agents involved (buyers and sellers) deploy scaling based strategies, and identify the Nash Equilibrium (NE) of the induced game. Researchers have used fictitious play-based convergence to equilibrium (e.g., (Shi et al. 2010)) in double auctions. However, such strategies are not useful in PDAs when the agents need to place bids in real-time for new auctions. In such settings, we believe scaling based strategies are easy to interpret and implement. The equilibrium analysis of non-linear or other complex forms are analytically difficult to compute; moreover may not be appealing to the real users of these markets.

We characterize NEs for One Buyer and One Seller (OBOS) and Two Buyer and One Seller (TBOS) analytically (Theorem 1 and 2). Given our assumption of scaling based bidding strategies and uniform type distributions, generic equilibrium analysis of double auctions, following ACPR, beyond these settings is challenging. To test our double auction strategies, we take the help of the PowerTAC simulation environment. PowerTAC is a simulation platform that
replicates crucial elements of the smart grid, where multiple distribution agencies (brokers) compete across markets to generate the most profit. Note that, the double auctions in PowerTAC; for that matter actually in electricity markets; are PDAs. In PDAs, the market clear multiple times, each after a specific time interval.

Now, if a buyer knows all the bids in a double auction, we argue that it is a best response for the buyer to bid as close as possible to the last clearing bid in order to procure the full required energy (Proposition 1). However, in reality, buyers never have access to such information. To address this incomplete information, we model the bidding process in PowerTAC PDAs as a Markov Decision Process (MDP), and solve it using dynamic programming and Last Clearing Price (LCP) prediction. Motivated by Power TAC’s fast response to solve it using dynamic programming and Last Clearing Price, it is referred to as “last clearing” since it is the last bid (ask) to be cleared by the clearing mechanism of the auction.

Definition 3. (Last Clearing Price (LCP)) Last Clearing Price (LCP) of bids (asks) refers to the limit-price of the Last Clearing Bid (Ask).

Definition 4. (The k-Double Auction) If a buyer and seller participate in a double auction, and if the sealed bid b by the buyer is higher than the sealed bid s by the seller, then CP is given by kb + (1 − k)s for some fixed k ∈ [0, 1].

Definition 5. (Average Clearing Price Rule (ACPR)) In a double auction, the clearing price and payment rule is ACPR if the clearing price is given by (b + s)/2 where b is the last executed bid, and s is the last executed ask. It is a special case of k-double auction, with k = 0.5.

Consider a game Γ = (N, (S_i)_{i∈N}, (u_i)_{i∈N}), where N = {1, 2, ..., n} is the set of players, S_i is the strategy set of the player i, and u_i : S_1 × S_2 × ... × S_n → R for i = 1, 2, ... n are utility functions.

Definition 6. (Best Response) Given a game Γ, the best response correspondence for player i is the mapping B_i : S_{i−1} → S_i defined by B_i(s_{i−1}) = {s_i ∈ S_i : u_i(s_i, s_{−i}) ≥ u_i(s'_i, s_{−i})}∀s'_i ∈ S_i. That is, given a profile s_{−i} of strategies of the other players, B_i(s_{−i}) gives the set of all best response strategies of player i.

Definition 7. (Nash Equilibrium) Given a game Γ, a strategy profile s^* = (s^*_1, s^*_2, ..., s^*_n) is said to be a Nash Equilibrium of Γ if, u_i(s^*_i, s^*_−i) ≥ u_i(s_i, s^*_−i)∀s_i ∈ S_i ∀i = 1, 2, ..., n. That is, each player’s Nash Equilibrium strategy is a best response to the Nash Equilibrium strategies of the other players.

Definition 8. (Markov Decision Process (MDP)) (Puterman 1994)) A Markov Decision Process (MDP) is a tuple given by M = (S, A, P, r, γ) where S is the set of states, A is the set of actions, P is the state transition probability function, where P(s'|s, a) = P(s_{t+1} = s'|s_t = s, a_t = a) is the probability that action a in state s at time t will lead to state s' at time t+1, r is the reward function, with r(s, a) denoting the reward obtained by taking action a in state s, and γ ∈ [0, 1] is the discount factor.

**Definitions & Background**

We first define all the required terms formally.

**Definition 1.** (Periodic Double Auction (PDA)) A type of auction, for buying and selling some resource, with multiple discrete clearing periods i.e. clearing after a specific time interval. Potential buyers submit their bids and possible sellers simultaneously submit their asks to an auctioneer. Then the auctioneer matches the bids and asks, and chooses some clearing price, denoted as CP, that clears the auction (Wurman, Walsh, and Wellman 1998). The allocation rule determines the quantity bought/sold by each buyer/seller, while the payment rule determines how much each buyer/seller pays/earns for buying/selling that quantity.

**PowerTAC** In this work, we focus more on smart grids. The Power Trading Agent Competition (PowerTAC) (Ketter, Collins, and Weerdt 2017) environment simulates a smart grid for approximately 60 days, where multiple brokers compete against each other across three markets - tariff, wholesale and balancing market - to generate the most profit. Each broker maintains a portfolio of consumers and producers, and buys and sells energy in the wholesale market. The broker with the highest bank balance at the end of the simulation, wins the game. We use the PowerTAC simulator to benchmark our bidding strategy.

The PowerTAC wholesale market employs PDAs for wholesale market energy trading. The clearing price and
A payment rule for the PowerTAC PDA, is given by ACPR. Three types of entities participate in these auctions - (1) Generating Companies (GenCos), (2) Miso Buyer, and (3) PowerTAC brokers. GenCos place only asks to sell energy, while the Meso Buyer places very low bid prices to buy energy. The PowerTAC brokers are free to place a bid or an ask depending on their requirement, or not place a bid at all.

The brokers can always participate in 24 auctions to trade energy, one auction for each of the next 24 timeslots. Each broker is notified about identity of other brokers participating in the PDAs at the beginning of the simulation. Each broker estimates its own energy requirement, and knows its own type. However, it does not know the types and requirements of the competing brokers. Every broker is allowed to submit an unlimited number of bids for each auction. After clearance, the clearing price and total cleared quantity of the auction is made public to all the brokers, while the last cleared bid or ask is not revealed. Additionally, each broker is privately notified about the cleared quantity and clearing price of any of its cleared bids/asks. The orderbook of the auction, which is the set of uncleared bids and asks without identity of the bidders, is also made public to all the brokers. If a broker fails to balance its retail demand portfolio after all the 24 auctions in the wholesale market, the balancing market automatically supplies the energy while charging the broker a balancing-price for its imbalance. The balancing-price is comparatively higher than the wholesale market price, and is meant to penalize the broker for having an imbalance. For more details about the PowerTAC simulation, we refer the reader to the Power TAC 2018 Game Specification (Ketter, Collins, and Weerdt 2017).

Related Work

Most bidding strategies for double auctions are designed for Continuous Double Auctions (CDAs) and would need to be modified for PDAs. Bidding strategies for PDAs, outside PowerTAC, are very limited. Wah, Hurd, and Wellman (2016) showed that in equilibrium, slow traders have higher welfare compared to fast traders in PDAs. As for bidding strategies for the PowerTAC wholesale market, AstonTAC (Kuate et al. 2013) uses Non-Homogeneous Hidden Markov Models (NHHMM) to predict energy demand and clearing price, which are then fed to an MDP to determine bid prices. TacTex (Urieli and Stone 2014; 2016a; 2016b) uses an MDP and dynamic programming based strategy derived from Tesauro and Bredin’s bidding strategy to predict bid prices, which is the motivation for our MDP-based strategy. Chowdhury 2016 predicts bid prices for the wholesale market PDAs using REPTree, Linear Regression and NN with weather data, with the former being is used in the SPOT (Chowdhury et al. 2017) broker. Chowdhury et al. 2018 use a Monte Carlo Tree Search (MCTS) based strategy coupled with a REPTree based price predictor (Chowdhury 2016) and heuristics, to determine optimal bid prices. AgentUDE (Özdemir and Unland 2015) uses an adaptive Q-learning based strategy in the wholesale market. None of these strategies are backed up by game theoretic analysis, where as our work is to build strategies derived from Nash Equilibrium.

Theoretical Approach and Proofs

In this section, we focus solely on the best response and Nash Equilibrium analysis of double auctions.

Nash Equilibrium analysis in single unit Double Auctions

Consider a single unit double auction, with the clearing price and payment rule given by ACPR. To find a generic Nash Equilibrium in this setting, we first try to simplify the double auction by restricting the number of buyers and sellers and their behavior. Upon doing so, we derive the following case-wise results.

One buyer and One Seller (OBOS)  Let’s assume that one buyer and one seller participate in the double auction, with their types as $\theta_B$ and $\theta_S$ respectively. We assume that both deploy scaling based strategies, i.e., a bid by a buyer is $b_B = \alpha_B \theta_B$ and an ask by the seller is $b_S = \alpha_S \theta_S$ where $\alpha_B$ and $\alpha_S$ are the scale factors by which the buyer and seller scale their true types while bidding, respectively. Motivated by the literature (Rothkopf 1980) (Vincent 1995) (Narahari 2014), we choose scale based bidding strategies for this Nash Equilibrium analysis, as compared to additive bidding strategies.

We assume $\theta_B \sim U[l_B, h_B]$ and $\theta_S \sim U[l_S, h_S]$ and this is common knowledge. We also assume Equation (1), which states that the buyer’s bid (seller’s ask) at any point will be less (higher) than or equal to the highest (lowest) possible seller’s ask (buyer’s bid).

$$\frac{\alpha_B}{\alpha_S} \theta_B \leq h_S, \quad \frac{\alpha_S}{\alpha_B} \theta_S \geq l_B$$

Thus, the utility of the buyer if its bid gets cleared, is denoted by the difference of true valuation and clearing price.

$$u_B = \int_{l_S}^{h_S} \left[ \theta_B - \left( \frac{\alpha_B \theta_B + \alpha_S \theta_S}{2} \right) \right] d\theta_S \quad (1)$$

Now assuming that the buyer decides to fix its $\alpha_B$ before even seeing its own type, then its utility is given by:

$$U_B = \int_{l_B}^{h_B} u_B \theta_B d\theta_B = \left[ \theta_B \left( 1 - \frac{\alpha_B}{\alpha_S} \theta_B - l_S \right) \right]_{l_B}^{h_B}$$

$$= \frac{\alpha_S}{4} \left[ \left( \frac{\alpha_B}{\alpha_S} \theta_B \right)^2 - l_B^2 \right] d\theta_B \quad (2)$$

$$= \left( \frac{h_B^3 - l_B^3}{3} \right) \left( \frac{\alpha_B}{\alpha_S} \theta_B - l_S \right) - l_S \left( 1 - \frac{\alpha_B}{2} \right) \left( \frac{h_B^2 - l_B^2}{2} \right)$$

$$+ \frac{\alpha_S}{4} l_B^2 \left( h_B - l_B \right)$$

Now, differentiating w.r.t. $\alpha_B$ and equating to 0 to find maxima:

$$\frac{\partial U_B}{\partial \alpha_B} = 0 \Rightarrow \left( \frac{h_B^2 - l_B^2}{3} \right) \left( \frac{1}{\alpha_S} - \frac{3 \alpha_B}{2 \alpha_S} \right) + l_S \left( \frac{h_B^2 - l_B^2}{4} \right) = 0$$

$$= \frac{2}{3} + \frac{\alpha_S}{\alpha_B} l_S \frac{h_B^2 - l_B^2}{2} \Rightarrow \alpha_B = \frac{4}{3} \frac{l_S x}{4 - l_S h_B x} \quad (4)$$
where $\frac{h_B^3 - l_B^3}{h_B^3 - l_B^3} = x$ and $\frac{h_S^3 - l_S^3}{h_S^3 - l_S^3} = y$.

Similarly, for the seller, assuming that the seller decides to fix its $\alpha_S$ before even seeing its own type, we get:

$$\alpha_S = \frac{4}{3} \left( \frac{2 + h_B y}{4 - l_S h_B y} \right) \tag{5}$$

Due to space constraints, we refer the reader to the extended version of our paper (Ghosh et al. 2019a) for the thorough derivation. Putting $l_S = l_B = 0$ and $h_S = h_B = 1$ in Equations (4) and (5), we get $\alpha_S = 1$ and $\alpha_B = \frac{2}{3}$. The above discussion is summarized as the following theorem.

**Theorem 1.** For a single unit double auction with ACPR, with only one buyer and one seller, whose true types are drawn from a $0 - 1$ uniform distribution, if they deploy scaling based bidding strategies $b_B$ and $b_S$ which satisfy Equation (1) and fix their scaling factors $\alpha_B$ and $\alpha_S$ before seeing their true types, then $\alpha_S = 1$ and $\alpha_B = \frac{2}{3}$ constitute a Nash Equilibrium.

**Two Buyers and One Seller (TBOS)** Let’s assume that two buyers $B_1$ and $B_2$, and one seller participate in the double auction, with types $\theta_{B1}, \theta_{B2}$ and $\theta_S$ respectively. We assume that all deploy scaling based bidding strategies, and both buyers have the same scaling factor $\alpha_B$. Thus, a bid by buyer $B_1$ is $b_{B_1} = \alpha_B \theta_{B1}$ and buyer $B_2$ is $b_{B_2} = \alpha_B \theta_{B2}$, while a bid by the seller is $b_S = \alpha_S \theta_S$. We also assume Equation (6), which states that the first buyer’s (seller’s) bid at any point will be less than or equal to the highest possible seller’s (buyer’s) bid.

$$\frac{\alpha_B}{\alpha_S} \theta_{B1} \leq h_S, \quad \frac{\alpha_S}{\alpha_B} \theta_S \leq b_B \tag{6}$$

First, we find the utility of the first buyer. We consider the following cases:

1. $b_{B1} \geq b_{B2} \Rightarrow \theta_{B1} \geq \theta_{B2}$ and $b_{B2} \geq b_S \Rightarrow \theta_S \leq \frac{\alpha_B}{\alpha_S} \theta_{B2}$
2. $b_{B1} \geq b_S \Rightarrow \theta_{B1} \geq \frac{\alpha_S}{\alpha_B} \theta_S$ and $b_{B2} \leq b_S \Rightarrow \theta_{B2} \leq \frac{\alpha_S}{\alpha_B} \theta_S$

Similar to OBOS, we assume that the first buyer decides to fix its $\alpha_B$ before even seeing its own type. We find its utility and differentiate it w.r.t $\alpha_B$ and equate it to 0 to find the maxima, and we get:

$$\left[ \frac{\alpha_B}{6 \alpha_S} (-2h_B^3 + 4l_B^3 + l_B h_B^2 - 2l_B h_B^2) + \frac{\alpha_S}{4 \alpha_B} (h_B^3 l_S^2 + l_B^2 h_S^2) \right]_{l_S = 1} = 0 \tag{7}$$

Similarly, for the seller, we consider the following 4 cases:

1. $b_{B1} \geq b_{B2} \Rightarrow \theta_{B1} \geq \theta_{B2}$ and $b_{B2} \geq b_S \Rightarrow \theta_{B2} \geq \frac{\alpha_S}{\alpha_B} \theta_S$
2. $b_{B2} \geq b_{B1} \Rightarrow \theta_{B2} \geq \theta_{B1}$ and $b_{B1} \geq b_S \Rightarrow \theta_{B1} \geq \frac{\alpha_S}{\alpha_B} \theta_S$
3. $b_{B1} \geq b_S \Rightarrow \theta_{B1} \geq \frac{\alpha_S}{\alpha_B} \theta_S$ and $b_{B2} \leq b_S \Rightarrow \theta_{B2} \leq \frac{\alpha_S}{\alpha_B} \theta_S$
4. $b_{B2} \geq b_S \Rightarrow \theta_{B2} \geq \frac{\alpha_S}{\alpha_B} \theta_S$ and $b_{B1} \leq b_S \Rightarrow \theta_{B1} \leq \frac{\alpha_S}{\alpha_B} \theta_S$

Again, similar to the buyer, we assume that the seller decides to fix its $\alpha_S$ before even seeing its own type. Then, we find its utility and differentiate w.r.t $\alpha_S$ and equate it to 0 to find maxima, and we get:

$$\left[ -\frac{\alpha_S^2}{4 \alpha_B} (h_S^4 - l_S^4) + \frac{\alpha_S}{\alpha_B} (h_S^3 l_B^2 - l_S h_B^2) - \frac{1}{6} (h_S^3 - l_S^3) (3l_B - h_B) \right]_{l_B = 1} = 0 \tag{8}$$

We refer the reader to our extended version of the paper (Ghosh et al. 2019a) for the thorough derivation. From Equation (7) and Equation (8), we get bi-variate cubic and quadratic equations in $\alpha_B$ and $\alpha_S$. Putting $l_S = l_B = 0$ and $h_S = h_B = 1$ in Equation (7) and Equation (8), we get $\alpha_S = \frac{1 \pm \sqrt{10}}{4} \approx 1.0406$ and $\alpha_B = \frac{2}{3}$. The above discussion can be summarized as the following theorem.

**Theorem 2.** For a single unit double auction with ACPR with two buyers and one seller, whose true types are drawn from a $0 - 1$ uniform distribution, if they deploy scaling based strategies $b_{B1}, b_{B2}$ and $b_S$, with buyers having the same scaling factor $\alpha_B$, which satisfy Equation (6) and fix their scaling factors $\alpha_B$ and $\alpha_S$ before seeing their true types, then $\alpha_S = \frac{1 \pm \sqrt{10}}{4}$ and $\alpha_B = \frac{2}{3}$ constitute a Nash Equilibrium.

As seen, with the increase in just one buyer, the complexity of the solution increases. It becomes increasingly difficult to extend and generalize the above results for a realistic market setting. Thus, moving forward, taking the PowerTAC wholesale market as testbed, we present a bidding strategy and experimentally show that it follows the theoretical results obtained in this section.

**Best Response analysis in multi-unit Double Auctions with complete information**

In practice, there are key differences between double auctions implemented in markets, and the theoretical results arrived above, stated as follows:

1. Quantity may be involved in the trading market auctions, which is not considered above.
2. The seller needs to use the same bidding strategy for one to achieve the above result, which may not be the case.

So, considering a multi-unit double auction with ACPR, where bids are of the form (quantity, price), we propose the best response if all the other bids are known to the bidder (i.e. complete information).

**Proposition 1.** When a buyer (seller) has complete information about the auction, and it desires to procure (sell) entire energy it bids (asks) for, it’s a best response to bid as close as possible to the last clearing bid (ask).

Since the proof is trivial, we refer the reader to the extended version of our paper (Ghosh et al. 2019a) for the same. Given the above proposition in a complete information setting, we further propose a MDP-based bidding strategy, which uses the past auction trends and statistics, to achieve the best response with incomplete information in the PowerTAC wholesale market.
MDPLCPBS: PowerTAC Wholesale Market Bidding Strategy

We introduce the MDP and LCP based Bidding Strategy (MDPLCPBS) for the PowerTAC wholesale market. The PowerTAC wholesale market accepts bids of the form (energy amount, limit-price). With respect to a broker, let the energy amount being sold be positive, while the energy amount being bought be negative. Meanwhile, let negative price indicate a broker is earning revenue, while positive price indicate it is paying or losing revenue. Thus, from the viewpoint of a broker, a buy order is seen to have a negative energy amount and a positive limit-price, while a sell order (termed as an ask), is seen to have a positive energy amount and a negative limit-price.

At timeslot \( t \), assuming a broker has a predicted demand profile \( D_t = \{d_{t+1}, d_{t+2}, \ldots, d_{t+24}\} \), where \( d_i \) is the predicted net demand at timeslot \( i \). Also, let \( P_t = \{p_{t+1}, p_{t+2}, \ldots, p_{t+24}\} \) denote the amount of energy already procured by past energy contracts, where \( p_i \) denotes the energy procured for timeslot \( i \). Thus, the remaining energy to be procured is given by \( E_t = \{e_{t+1}, e_{t+2}, \ldots, e_{t+24}\} \), where \( e_i = d_i - p_i \) is the net energy left to be procured for timeslot \( i \). The bidding strategy, MDPLCPBS, to procure the aforementioned energy requirements, comprises of three major submodules - (i) Limit Price Predictor, (ii) Quantity Predictor, and (iii) Last Cleared Price Predictor.

Limit Price Predictor (LPP)

At any given timeslot \( t \), the predictor computes 24 limit-prices for 24 simultaneous PDAs in the PowerTAC wholesale market. Motivated by (Tesauro and Bredin 2002) and (Urieli and Stone 2014), the Limit Price Predictor uses the following MDP to place optimal limit-prices for bids:

1. **States**: \( s \in S = \{0, 1, \ldots, 24, success\} \), \( s_0 := 24 \)
2. **Actions**: limit-price \( \in \mathbb{R} \)
3. **Transition**: The same state transition from (Urieli and Stone 2014) is used. A state \( s \in \{1, \ldots, 24\} \) transitions to one of two states. If a bid is partially or fully cleared, it transitions to the terminal state success. Otherwise, a state \( s \) transitions to state \( s - 1 \). The clearing (i.e. transition) probability \( p_{\text{cleared}}(s, \text{limit-price}) \) is initially unknown and is determined by Equation (10).
4. **Reward**: At any state \( s \in \{1, \ldots, 24\} \), the reward is 0. At terminal state \( s = 0 \), the reward is the negative of the balancing-price per unit energy. At terminal state \( s = success \), the reward is the negative of the limit-price of the cleared bid. Since we take the price to be positive for bids and negative for asks, maximizing reward results in minimizing costs.
5. **Terminal States**: \( \{0, success\} \)

We solve the above MDP using a sequential bidding strategy, that computes the optimal bid limit-price that minimizes the expected procurement cost per unit energy. It uses the balancing-price as the expected price at state \( s = 0 \), and recursively minimizes the expected cost by using the probability of clearance, \( p_{\text{cleared}}(s, \text{limit-price}) \). This solution is summarized as a value function, stated as follows:

\[
V(s) = \begin{cases} 
\text{balancing-price,} & \text{if } s = 0 \\
\min_{\text{limit-price}} \left\{ p_{\text{cleared}} \times \text{limit-price} \right\} & \text{if } s \in [1, 24] \\
+(1 - p_{\text{cleared}}) \times V(s - 1), & \text{if } s = 24
\end{cases}
\]

(9)

Given that the balancing-price and the \( p_{\text{cleared}} \) values are different for bids and asks, we maintain two separate instances of the MDP, and solve them independently.

The value function in Equation (9) is solved recursively using dynamic programming. However, before doing so, the balancing-price and the transition function \( p_{\text{cleared}}(s, \text{limit-price}) \) need to be estimated, as they are both initially unknown. The balancing-price is estimated by averaging the balancing-prices across past timeslots. On the other hand, the clearing probability, \( p_{\text{cleared}}(s, \text{limit-price}) \), is computed using past auction statistics as:

\[
p_{\text{cleared}}(s, \text{limit-price}) = \frac{\sum_{\text{auction} \in \text{ac}} \text{ac.LCP}\times\text{limit-price}\times\text{ac.cleared-amount}}{\sum_{\text{auction} \in \text{ac}} \text{ac.cleared-amount}}
\]

(10)

where \( \text{auction} \) is the set of all past auctions in the state \( s \), and LCP is the Last Clearing Price, which is estimated by the Last Cleared Price Predictor. The auction statistics for each state \( s \) are re-used in the future for estimating \( p_{\text{cleared}} \), as we iterate over the same sequence of states \( S \) during the bidding process.

Quantity Predictor (QP)

The Quantity Predictor is primarily responsible for distributing the demand for a target timeslot across all the 24 auctions, in order to further reduce overall energy cost. The idea is to buy more and sell less at cheaper prices, and vice-versa. It essentially breaks down the demand for a target timeslot across all the 24 auctions, in order to further reduce overall energy cost. The energy quantity to bid/ask, for each state \( s \) at timeslot \( t \), is given by:

\[
q(s) = \begin{cases} 
\frac{e_{t+s}}{\sum_{j=1}^{s} \text{limit-price}[j]}, & \text{if } e_{t+s} > 0 \\
\frac{e_{t+s}}{\sum_{j=s}^{24} \text{limit-price}[j]}, & \text{if } e_{t+s} < 0 \\
0, & \text{if } e_{t+s} = 0
\end{cases}
\]

(11)

where \( s \in \{1, \ldots, 24\} \), limit-price\( [s] \) is the limit-price for state \( s \) determined by the Limit Price Predictor. The first case in Equation (11) refers to the situation where energy needs to sold, so the bid quantity is directly proportional to the predicted limit-price of that auction - essentially selling more energy at higher price. On the other hand, the second case occurs when the energy needs to be procured. So, the bid quantity is set to be inversely proportional to the predicted limit-price i.e. buying more energy at cheaper price. Thus, the final bid is of the form \((q(s), \text{limit-price}[s])\).
Last Cleared Price Predictor (LCPP)

First, one has to note that, in any auction, the LCP is greater than or equal to CP. Mostly, \( LCP \geq CP \), as \( P(LCP = CP) = 0 \), i.e. LCP equal to CP is a probability zero event. In PowerTAC, the LCP is not known to any broker. In essence, one can place better bids if the LCP for each auction is known, as they can bid higher than a predicted LCP to become the last bid, and achieve best response according to Proposition 1. The Last Cleared Price Predictor essentially tries to determine the LCP for bids and asks for all executed auctions. It does so by probing the auctions with a set of dummy orders, which have the minimum tradeable energy as quantity (0.01 MwH), and limit-prices equally spaced in the range \( \beta \times \text{limit-price, balancing-price} \). After execution, the LCP for bids for an auction in state \( s \) is determined by:

\[
LCP(s) = \min(\text{dummy-bids}_{\text{cleared}}, \text{limit-price}[s]_{\text{cleared}})
\]

where \( \text{dummy-bids}_{\text{cleared}} \) is the set of bid prices of all dummy bids which got cleared in the state \( s \), and \( \text{limit-price}[s]_{\text{cleared}} \) is the limit-price for the cleared final bid made in state \( s \) (taken to be infinity if final bid did not clear or does not exist). Similarly, the LCP for asks is given as:

\[
LCP(s) = \max(\text{dummy-asks}_{\text{cleared}}, \text{limit-price}[s]_{\text{cleared}})
\]

where \( \text{dummy-asks}_{\text{cleared}} \) is the set of ask prices of all dummy asks which got cleared in the state \( s \), and \( \text{limit-price}[s]_{\text{cleared}} \) is the limit-price for the cleared final ask made in state \( s \) (taken to be infinity if final ask did not clear or does not exist). These LCP values are then used to update the clearing probability \( p_{\text{cleared}} \) in Equation (10). Algorithm 1 summarizes MDPLCPBS, which is executed every timeslot. It takes the energy requirement for the 24 auctions as input. First it collects the market statistics, which includes the LCP estimate and clearing amount from previous timeslots, and the balancing price (line 2). If the number of data points is suitable enough (taken to be 24 in our implementation), it proceeds to solve the MDP and generates a set of prices to bid (line 4). Using these set of prices, and the energy requirements, it generates a set of quantities to bid (line 5). If data points are not enough, the bidding policy

\[
\text{Algorithm 1 MDPLCPBS}
\]

1: procedure MDPLCPBS(energyReq[1..24])
2: \hspace{1em} marketData[0..24] \leftarrow \text{getMarketStatistics}()
3: \hspace{2em} if EnoughDataPoints(marketData) then
4: \hspace{3em} \hspace{1em} bidPrices[1..24] \leftarrow \text{SolveMDP(marketData)}
5: \hspace{3em} \hspace{1em} bidQty[1..24] \leftarrow \text{SpreadQty(energyReq, bidPrices)}
6: \hspace{2em} else
7: \hspace{3em} \hspace{1em} bidPrices[1..24] \leftarrow \text{SampleBiddingPolicy()}
8: \hspace{3em} \hspace{1em} bidQty[1..24] \leftarrow \text{energyReq[1..24]}
9: \hspace{2em} end if
10: \hspace{1em} sendBids(bidPrices, bidQty)
11: \hspace{1em} sendDummyBids(bidPrices, marketData)
12: end procedure

given in the PowerTAC sample-broker is used to determine the bid prices (line 7), and the bid quantities are set as the full energy requirements (line 8). Using the determined bid prices and quantities, we place the actual bids (line 10), and a set of dummy bids in the market (line 11). The time complexity of Algorithm 1 comes out to be in the order of the number of past market data points.

Experimental Analysis

We first analyze if our proposed bidding strategy, MDPLCPBS, follows the Nash Equilibrium derived above, and then benchmark it against the baseline and competing state-of-the-art strategies.

Validation Experiments

We take the Power TAC simulator and isolate the wholesale market, and remove all market simulator participants (GenCos, internal buyers) from the market. We test the one buyer one seller (OBOS) scenario by deploying only two agents in the isolated wholesale market - a buyer and a seller. These agents have a fixed energy demand that they need to buy (sell) from the market. In these experiments, we set the energy demand to be the previous slot’s tariff market net demand, which both the buyer and seller are notified about. We draw \( \theta_B \sim U[40,80] \) and \( \theta_S \sim U[40,80] \), and compute the theoretical scale factors using Equation (4) and Equation (5). We run two batches of experiments, with 30 games in each set of the batch, for 5 sets per batch. During each batch, one of the agents has a fixed scaling based bidding strategy, while the other uses MDPLCPBS.

In the first (second) batch, we draw the seller’s (buyer’s) valuation \( \theta_S \sim U[40,80] \) (\( \theta_B \sim U[40,80] \)), and apply a fixed scale factor within \( \pm0.1 \) of the theoretical value. The buyer (seller) generates its valuation \( \theta_B \sim U[40,80] \) (\( \theta_B \sim U[40,80] \)), and uses this valuation as the balancing-price in MDPLCPBS to generate bids.

The experimental average scale factor and standard deviation, for cleared bids, for the buyer and the seller in the two batches of experiments are documented in Table 1. The table demonstrates that in a one buyer and one seller setting, MDPLCPBS approaches the Nash Equilibrium characterized for a single unit OBOS double auction. The values demonstrate that as the fixed scale factor for the seller is increased, the buyer’s scale factor increases slowly. On the contrary, when the fixed scale factor for the buyer is increased, the seller’s scale factor increases rapidly.

Benchmarks

We isolate the PowerTAC wholesale market from the full PowerTAC simulator while keeping the market simulator participants (GenCos, internal buyers) and weather simulator, and benchmark the performance of MDPLCPBS. The following agents/brokers are used in these benchmarks:

- Zero Intelligence (ZI): The ZI agent (Gode and Sunder 1993) uses a randomized bid strategy and ignores the market state. It generates random order prices, ignoring the state of the market. In our experiments, we derive its bids from a uniform distribution with mean \( \mu \) and a standard
deviation of $10. The mean $\mu$ taken from the limit price predicted by the MDP in TacTex (Urieli and Stone 2014). The broker places one bid per auction, and the remaining required energy as the bid quantity. It continues to do the same for all the 24 bidding opportunities, or until the required energy is procured.

- Zero Intelligence Plus (ZIP): The ZIP agent (Tesauro and Das 2001) maintains a scalar variable $m$ denoting its desired profit margin, and it combines this with a unit’s limit price to compute a bid price $p$. For each failed trade, the price is adjusted by small increments to beat the failed bid price $p$. In our experiments, the initial limit price value $\mu$ is determined from the limit price predicted by the MDP in TacTex. The profit margin $m$ is set to $1/\%$ of $\mu$, resulting in the initial bid price to be $p = \mu \times 1.01$. If the bid fails, the next bid price is incremented by 10% of $\mu$. Then, the new bid price is given by $p = \mu \times 1.11$.

- TacTex: The TacTex (Urieli and Stone 2014) agent uses an MDP based model and dynamic programming to determine limit-prices for bids. The algorithm described in the paper was implemented and used in our experiments.

- MCTS: The MCTS (Chowdhury et al. 2018) agent uses a Monte Carlo Tree Search (MCTS) coupled with heuristics on top of the limit price derived from a REPTree based limit price predictor, to determine the optimal bid price. In our experiments, we used the MCTS-dyn-C2 version with 10000 iterations, which is shown to be the best performing variation of the MCTS bidding strategy.

For a timeslot $t + 24$ in the future, having 24 bidding opportunities in timeslots $\{t, t + 1, \ldots, t + 23\}$, the energy to be procured is set to be same across all the brokers. This energy amount for $t + 24$ is determined as some fraction of the net demand in timeslot $t$ in the PowerTAC simulation tariff market. Four sets of 10 games each are simulated, with each set having a different fraction of the net demand to be procured. The fraction set is given by $\{0.25, 0.5, 0.75, 1\}$.

Figure 1 shows the net cost of all the agents across the four sets of games. In each case, MDPLCPBS outperforms ZIP, ZIP and TacTex on a consistent basis, while losing out to MCTS. While ZIP may seem to perform reasonably well in some cases, it can be countered easily in strategic settings, like in single-shot single-unit auction setting, whereas MDPLCPBS follows the equilibrium. It is also to be noted that, while MCTS uses tailored heuristics, MDPLCPBS is derived from the game theoretic analysis of single shot double auction (Proposition 1). We leave the game theoretic analysis of MCTS for future work.

Table 2 summarizes the weighted relative error rates (weighed by cleared quantity) in predicting the LCP. 82% of the total cleared energy for a future timeslot is cleared in the first auction itself, with 91% of the total being cleared in the first five. Since the cleared energy of the other states is extremely low, their corresponding predicted LCPs have less impact on the $p_{\text{clear}}$ calculation. Thus, we focus on the error rates of the states corresponding to the first five auctions for a target timeslot. We see that MDPLCPBS has a 10% and 15% error rate in the LCP prediction for the first auction of a timeslot (state 24), for 100% and 50% of the market demand as requirement. Moreover, the corresponding weighted average relative error across all states, comes out to be 12% and 18% respectively. The error rates increase as requirement decrease, as there are more low quantity bids by brokers and Miso buyer’s low bids often set the LCP. Thus, MDPLCPBS
predicts the LCP with minimal error, during auctions where most of the energy gets traded.

Conclusion
In this paper, we first analytically characterized Nash Equilibrium strategies for a single unit double auction with the clearing price and payment rule as ACPR, for OBOS, and TBOS with scale based bidding strategies. We also proposed the best response in a complete information setting in a multi-unit double auction with ACPR. Based on these formulations, we presented MDPLCPBS, a bidding strategy for PDAs. Furthermore, we experimentally validated that MDPLCPBS achieves the Nash Equilibrium derived for single unit double auction with ACPR for OBOS. Finally, we benchmarked MDPLCPBS against the baseline and competing state-of-the-art strategies, and showed that it outperforms most of them consistently. Simultaneously, we showed that it predicts the LCP with minimal error.

References