Optimizing Prosumer Policies in Periodic Double Auctions Inspired by Equilibrium Analysis

Bharat Manvi[1] , **Sanjay Chandlekar**[1,2] and **Easwar Subramanian**[1]
[1]TCS Research
[2]IIIT Hyderabad
{bharat.manvi, easwar.subramanian}@tcs.com, sanjay.chandlekar@research.iiit.ac.in

Abstract

We consider a periodic double auction (PDA) wherein the main participants are wholesale suppliers and brokers representing retailers. The suppliers are represented by a composite supply curve and the brokers are represented by individual bids. Additionally, the brokers can participate in small-scale selling by placing individual asks; hence, they act as prosumers. Specifically, in a PDA, the prosumers who are net buyers have multiple opportunities to buy or sell multiple units of a commodity with the aim of minimizing the cost of buying across multiple rounds of the PDA. Formulating optimal bidding strategies for such a PDA setting involves planning across current and future rounds while considering the bidding strategies of other agents. In this work, we propose Markov perfect Nash equilibrium (MPNE) policies for a setup where multiple prosumers with knowledge of the composite supply curve compete to procure commodities. Thereafter, the MPNE policies are used to develop an algorithm called MPNE-BBS for the case wherein the prosumers need to reconstruct an approximate composite supply curve using past auction information. The efficacy of the proposed algorithm is demonstrated on the PowerTAC wholesale market simulator against several baselines and state-of-the-art bidding policies.

1 Introduction

Auctions play a crucial role in the real world, serving as dynamic marketplaces wherein the value of a commodity gets determined by supply and demand in real-time [Wikipedia contributors, 2023]. Double auctions are the most prominent type of auction, with a trade volume of more than a trillion dollars daily in stock exchanges [Parsons *et al.*, 2011] and energy markets [Ketter *et al.*, 2020]. Such auctions involve bids from the buyers and asks from the sellers; these bids and asks are submitted as a tuple comprising the desired unit price for the commodity and the intended number of units for procurement (or sale). Subsequently, the auction mechanism determines each participant's clearing price and quantity.

A periodic double auction (PDA) is a setup wherein buyers and sellers engage in a (finite) sequence of auctions to trade certain units of a commodity. For example, a buyer with the intention of procuring certain units of a commodity will engage in multiple (but finite) trades with the seller to satisfy her desired procurement target. These types of auctions are prevalent in energy markets wherein a power generating company (GenCo) that sells energy in bulk is the prominent seller and retail brokers are the buyers. In addition, PDAs are also used to model the call auctions in financial markets [Constantinides and Cartlidge, 2021]. The PDA setup allows buyers and sellers to trade energy periodically until a few hours before delivery. As players need to participate in a series of auctions, devising a bidding strategy that caters to current as well as future auctions is a challenging problem. Considering that energy trades in a smart grid setup enable trades worth 1000 TWh of energy, resulting in daily transactions of more than 3 billion dollars just in Europe [Disbrey and Zeier, 2020], it is prudent to design efficient bidding strategies on behalf of a market participant to bring in cost optimization and ecosystem efficiency. To this end, we model PDAs as a finite horizon Markov game and propose equilibrium solutions to aid in devising efficient bidding strategies.

Developing efficient bidding strategies depends mainly on clearing and payment rules of the auction mechanism. Although equilibrium solutions for double auctions have been studied under various payment rules [Wilson, 1992] including some existence results for average clearing price rule [Satterthwaite and Williams, 1989], multi-buyer [Ghosh *et al.*, 2020] and multi-item settings [Chandlekar *et al.*, 2022b], all these works involve simple double auctions and hence may not be applicable for PDAs as devising bidding strategies in PDA involves sequential decision-making. Some recent works do model PDAs as a Markov game and use techniques from reinforcement learning such as multi-agent Q learning [Rashedi *et al.*, 2016], multi-agent deep Q network [Ghasemi *et al.*, 2020], deep deterministic policy gradients [Du *et al.*, 2021; Chandlekar *et al.*, 2022b]. Nevertheless, many of these works involve single-sided auctions[1] and involve numerical simulations rather than analytical solutions. A recent work [Manvi and Subramanian, 2023] did propose an equilibrium solution by modelling the PDA as a complete

[1]Auctions wherein only buy or sell bids are placed

information Markov game. However, the work is limited to cases where the clearing mechanism is the average clearing price rule (ACPR) and supply by wholesale suppliers is adequate. Furthermore, the players involved in the auction were only buyers, not the prosumers. Moreover, the said work does not provide any policy for a realistic auction setting, which involves incomplete information.

Our work overcomes the abovementioned limitations by providing equilibrium strategies for prosumers (not just buyers) for any general uniform clearing rule satisfying certain properties. We then propose a novel bidding strategy MPNE-BBS inspired by this analytical equilibrium solution that approximately reconstructs the supply curve from past auction information to place bids in the market. Our distinct contributions are as follows,

1. We consider a Markov game framework wherein the retail market brokers (net buyers or prosumers) can also participate in small-scale selling, enabling retailers with solar panels and EVs to sell energy along with buying.
2. We define the properties of the uniform clearing rule of the double auction, which encompasses various clearing mechanisms prevailing in diverse auction setups across different geographies.
3. We propose novel Markov perfect Nash equilibrium (MPNE) solutions when prosumers compete to procure the required commodities. The proposed solutions are valid for any uniform clearing mechanism of the PDA with the above mentioned properties.
4. Based on the solution concept derived for the complete information setting, we propose a novel bidding strategy, MPNE-BBS, to work in the *incomplete information* setup where the supply curve and information regarding the demand requirement of other buyers are not known.
5. We then demonstrate the efficacy of the MPNE-BBS algorithm against several baseline and state-of-the-art bidding strategies deployed in the close-to-real-world energy market simulator PowerTAC.

2 Related Work

Some of the early work in the double auction setting was devoted to obtaining equilibrium solutions for single buyer and single seller with a uniform distribution of valuations [Chatterjee and Samuelson, 1983]. [Satterthwaite and Williams, 1989] attempt to find non-trivial equilibria and show the existence of a multiplicity of equilibria for the k-double auction for a generic class of market participants' valuations. They propose the equilibrium strategies in the form of differential equations and then examine the efficiency of the proven equilibrium. A work [Krausz and Rieder, 1997] considers analytical solutions for a two-player zero-sum Markov game of incomplete information; whereas this work considers the multi-player general-sum game. [Vetsikas, 2014] proposes equilibrium strategies for multi-unit sealed bid auction for m^{th} and $(m+1)^{th}$ price sealed bid auction, which differ from k-double auctions in clearing rules.

PowerTAC [Ketter *et al.*, 2020] is a widely adopted platform to validate bidding strategies in PDAs and has a vast literature on bidding strategies. Various works have proposed Markov Decision Process (MDP) based strategies; for instance, [Urieli and Stone, 2014]'s MDP-based strategy was inspired by Tesauro and Bredin's bidding strategy [Tesauro and Bredin, 2002], which they solve using dynamic programming. Urieli and Stone's strategy was improved upon by [Ghosh *et al.*, 2020], where the authors also provide equilibrium analysis for single-item single-shot double auctions. [Chandlekar *et al.*, 2022b] present an analytical equilibrium solution for single-shot multi-unit auctions to design a DDPG-based bidding strategy. [Kuate *et al.*, 2013] also proposed an MDP-based bidding strategy to determine the bid quantity and use of Non-Homogeneous Hidden Markov Models (NHHMM) to determine bid prices. Additionally, some works have also adopted a Monte Carlo Tree Search (MCTS) framework to device bidding strategies for PDAs [Chowdhury *et al.*, 2018; Orfanoudakis *et al.*, 2021].

At best, all the strategies mentioned above involve equilibrium analysis of single-shot double auctions. Hence, the analysis may not be readily extended to the PDA setting which is a multi-shot auction and the current work bridges this gap.

3 Market Clearing Mechanism

We begin by describing the market clearing mechanism of a double auction. Consider a group of N prosumers who want to procure multiple units of a commodity from a group of sellers by participating in a PDA having H rounds. At any round $h \in [H]$[2] of the PDA, a prosumer $b \in [N]$ has $Q_+^{b,h}$ and $Q_-^{b,h}$ units of a commodity to buy and sell, respectively. We let $Q_+^{b,h}, Q_-^{b,h}$ to be unique $\forall b \in [N]$ with $Q_+^{b,h} > Q_-^{b,h}$. The set of buy bids is denoted as $\mathcal{B}_+^{b,h}$ with $B_+^{b,h}$ elements and the sell bids[3] are denoted by $\mathcal{B}_-^{b,h}$ with $B_-^{b,h}$ elements. The buy bids are the pair of price and quantity denoted as $(p_{i,+}^{b,h}, q_{i,+}^{b,h}),\ i \in [B_+^{b,h}]$, where price and quantity are bounded by $p_{\max}$ and $Q_+^{b,h}$ respectively. Similarly, the sell bids are denoted as $(p_{n,-}^{b,h}, q_{n,-}^{b,h}),\ n \in [B_-^{b,h}]$, with $p_{\max}$ and $Q_-^{b,h}$ as price and quantity upper bounds. As a result, the total outstanding demand in round $h \in [H]$ from all N prosumers is denoted as $Q^{\mathcal{D},h} = \sum_{b\in[N]} Q_+^{b,h}$.

The wholesale sellers, on the other hand, are represented by a consolidated supply curve with L^h asks expressed as $\mathcal{L}^h = \{(p_i^h, q_i^h) \mid i \in [L^h]\}$, with $p_i^h \in [0, p_{\max}]$ and $q_i^h \in [0, q_{\max}]$ as the price and quantity components of the i^{th} ask (p_i^h, q_i^h), with $p_{\max}$ and $q_{\max}$ as suitable upper bounds. Accordingly, the total supply provided by wholesale sellers, at round h, is denoted as $Q^{\mathcal{S},h} = \sum_{m\in[L^h]} q_m^h$. Now, the overall supply at round h of the PDA is given by $\hat{Q}^{\mathcal{S},h} = Q^{\mathcal{S},h} + Q_-^{\mathcal{S},h}$, where $Q_-^{\mathcal{S},h} = \sum_{b\in[N]} Q_-^{b,h}$ is the supply from the prosumers. Finally, at the round h, the combined asks and sell bids of the wholesale suppliers and the brokers is expressed as $\hat{\mathcal{L}}^h = \{(\hat{p}_i^h, \hat{q}_i^h) \mid i \in [L^h + B_-^h]\}$.

The market regulator at each round h, collects the elements of $\hat{\mathcal{L}}^h$ and $\mathcal{B}_+^h$ and uses a *clearing rule* to produce a clearing

[2] For any integer K, we denote $[K]$ as the set $\{1, \dots, K\}$.

[3] Sell bids are prosumer's asks.

price(s) and cleared quantities. In this work, we focus on the *uniform clearing rule*, where all the cleared bids and asks have the same clearing price. The cleared price for round h is called the market clearing price (denoted as λ^h) and the total quantity cleared for all the B_+^h bids at round h is denoted as Q^h. To further elaborate the clearing process, we assume, without loss of generality, that the elements of the set $\hat{\mathcal{L}}^h$ ($\mathcal{B}_+^h$) are sorted in increasing (decreasing) order of the price component. If the price components of the elements are equal, then $\hat{\mathcal{L}}^h$ and $\mathcal{B}_+^h$ are sorted in decreasing order of the quantity component. If both price and quantity components of certain elements are equal, then the ordering between them is arbitrary. More concretely, we consider the uniform clearing rules, which have properties defined in Definition 1.

Definition 1. *Given*

- $(\hat{p}_d^h, \hat{q}_d^h) \in \hat{\mathcal{L}}^h$ *as the last cleared ask and* $(p_{l,+}^{b,h}, q_{l,+}^{b,h}) \in \mathcal{B}_+^h$ *as the last cleared buy bid at round h.*
- $\alpha_i^{b,h}$ *with* $i \in [B_+^{b,h}]$ *and* $\beta_k^{b,h}$ *with* $k \in [B_-^{b,h}]$ *as the cleared buy and sell bid quantities of a prosumer* $b \in [N]$ *at round h.*

The ***properties*** *of the considered uniform clearing rules are defined as follows.*

- *The uniform clearing price* λ^h *at round h is a scalar value that lies in the interval* $[\hat{p}_d^h, p_{l,+}^{b,h}]$.
- *The total market cleared quantity at round h is,*
$$Q^h = \min\left\{\sum\nolimits_{j=1}^{d} \hat{q}_j^h, \sum\nolimits_{i=1}^{l} q_{i,+}^{b,h}\right\}.$$
- *Buy bids that are higher than the last cleared buy bid are fully cleared. That is ,* $\alpha_m^{b,h} = q_{m,+}^{b,h}$, *if* $p_{m,+}^{b,h} > p_{l,+}^{b,h}$.
- *Buy bids that are lower than the last cleared buy bid are not cleared. That is,* $\alpha_m^{b,h} = 0$, *if* $p_{m,+}^{b,h} < p_{l,+}^{b,h}$.
- *Buy bids that are equal to the last cleared buy bid are cleared as* $\alpha_m^{b,h} = \frac{1}{|B_{b=l}|}\left(Q^h - \sum_{j=1}^{|B_{b>l}|} q_{j,+}^{b,h}\right)$, *where* $|B_{b=l}|$ *and* $|B_{b>l}|$ *denote the number of bids that are equal and higher than the last cleared bid, respectively.*
- *Conversely, the asks and sell bids from* $\hat{\mathcal{L}}^h$ *are cleared such that the cheaper asks are given the higher priority.*

In practice, regulators deploy a variety of clearing rules that satisfy the properties defined in Definition 1. Popular examples include *merit order dispatch* [Taylor, 2015] and k-double auctions [Angaphiwatchawal *et al.*, 2021]. The merit order dispatch is a mechanism where the market operator clears the market by maximizing the area between supply (asks) and demand curves (bids). Specifically, market clearing is posed as an optimization problem, and the (primal and dual) solutions to this optimization give the cleared quantities and clearing price. The clearing price for the k-double auction (where $k \in [0\ ; 1]$) is given by $\lambda^h = k \cdot \hat{p}_d^h + (k-1) \cdot p_{l,+}^{b,h}$. Furthermore, the average clearing price mechanism (ACPR) is a special case of k-Double auction with $k = 0.5$, and its clearing price is given as the average of the last cleared ask and bid. For further details about the clearing mechanism please refer to [Manvi *et al.*, 2024].

4 The Markov Game Structure

Consider a Markov game [Zhang *et al.*, 2023] with a finite horizon to model the PDA with N prosumers. Specifically, let $\mathcal{G} = \langle N, S, A, C, P, H\rangle$ denote the Markov game with N as the number of players (prosumers), S as the state space, A as the joint action space, $C = \{C^{b,h} \mid b \in [N],\ h \in [H]\}$ as the cost functions for the players, with $C^{b,h}$ denoting the cost function for the player b at round h, $P : S \times A \to S$ as the transition probability and H as the length of the horizon.

More concretely, the state space S is defined as the set of wholesale suppliers' asks $\mathcal{L}^h$, brokers' demand Q_+^h and brokers' supply Q_-^h and thus S is $\{\mathcal{L}^h, Q_+^h, Q_-^h\}$. Here, Q_+^h and Q_-^h are sets of brokers' demand $Q_+^{b,h}$ and supply $Q_-^{b,h}$ for all $b \in [N]$. Note that since we are working with the *complete information setting*, the wholesale suppliers' asks, brokers' demand and brokers' supply are known to all the brokers; hence, the transition of the states is deterministic. The joint action space $A = \times_{b\in[N]} A^b$ is a product space of all the prosumer's action space A^b, where $A^b = \cup_{h\in[H]}(\mathcal{B}_+^{b,h} \times \mathcal{B}_-^{b,h})$ is the product space of the all the buy bids $\mathcal{B}_+^{b,h}$ and sell bids $\mathcal{B}_-^{b,h}$ over $h \in [H]$. The cost function $C^{b,h} : S \times A^b \to \mathbb{R}$ for a player b is given by

$$C^{b,h}(s^h, a^h) = \begin{cases} \Big[\sum_{n\in[B_+^{b,h}]} \lambda^h \cdot \alpha_n^{b,h} \\ \quad - \sum_{m\in[B_-^{b,h}]} \lambda^h \cdot \beta_m^{b,h} \Big] & 0 \le h \le H \\ \Upsilon \times Q^{b,h} & h = H+1, \end{cases}$$

where $a^h = (a^{b,h}, a^{-b,h}) \in A$ denote the joint action, $a^{-b,h}$ denote the actions taken by players other than b, $\alpha_n^{b,h}$ are cleared quantities for player b's buy bids and $\beta_m^{b,h}$ are the cleared sell bids. The constant $\Upsilon \ge 0$ is the *balancing price* required to buy the quantity $Q_+^{b,H+1}$ outside of the PDA at time $H+1$.

A state s^h at round h, when a joint action $a^h \in A$ is taken, transitions to a state $s^{h+1} = \{\mathcal{L}^{h+1}, Q_+^{h+1}, Q_-^{h+1}\}$ at round $h+1$. Here, $\mathcal{L}^{h+1}$ represents the uncleared asks at round $h+1$ and $\{Q_+^{h+1}, Q_-^{h+1}\}$ are the updated demand and supply respectively of the prosumers after accounting for the cleared quantities at round h. In addition, the sequence of transitions $s^1, a^1, s^2, a^2, \ldots, s^H, a^H, s^{H+1}$, which starts at s^1 and ends at s^{H+1}, is denoted as a trajectory τ of the Markov game. Here, the trajectory τ induces a sequence of costs $C^{b,1}, \ldots, C^{b,H}, C^{b,H+1}$ for each player $b \in [N]$. A player b will choose an action $a^{b,h} \in A^b$ at each of the state s^h she visits and the collection of actions taken at each round h is denoted by a Markov policy $\pi^b = \{\pi^{b,h} : S \to A^b \mid h \in [H]\}$. We further let $\pi = (\pi^b, \pi^{-b})$ as the joint policy that includes player b's policy π^b along with the policies of players except b, denoted as π^{-b}. Moreover, we let Π^b denote the policy space of player b and $\Pi = \times_{b\in[N]}\Pi^b$ denote the joint policy space. Importantly, to capture the cost of acquisition, we define a value function $V_\pi^h : S \to \mathbb{R}$ of a joint policy $\pi \in \Pi$ at round h in Equation (1).

$$V_\pi^h(s) = \sum\nolimits_{r=h}^{H+1} C^{b,r}(s^r, a^{b,r}, a^{-b,r}). \qquad (1)$$

More specifically, the value function is given by

$$V_\pi^h(s) = \sum_{r=h}^{H} \lambda^r \cdot \left(\sum_{n \in [B_+^{b,r}]} \alpha_n^{b,r} - \sum_{m \in [B_-^{b,r}]} \beta_m^{b,r} \right) + \Upsilon \cdot Q_+^{b,H+1}.$$

For the game $\mathcal{G}$, we define MPNE [Yang and Wang, 2020] using the value function in Definition 2.

Definition 2. *Given the game $\mathcal{G}$ of H horizon with N players, a joint policy $\pi_* = (\pi_*^b, \pi_*^{-b})$ is an MPNE if $\forall b \in [N]$, $\forall s \in S$, $\forall h \in [H]$ and $\forall \pi^b : S \to A^b$, we have*

$$V^h_{\pi_*^b, \pi_*^{-b}}(s) \le V^h_{\pi^b, \pi_*^{-b}}(s). \tag{2}$$

Having described the Markov game framework, we now proceed to develop MPNE solutions in the next section.

5 Equilibria of the Markov Game

We begin by focusing on the case where the prosumers $b \in [N]$ are restricted to place at most one buy bid and one sell bid. In addition, first, we consider that the supply from the bulk sellers is enough to satisfy the demand, that is $Q^{\mathcal{S},h} \ge Q^{\mathcal{D},h}$. Later, we consider the inadequate supply case, where the equilibrium solution differs from the adequate supply case. Note that it is essential to consider the inadequate supply case since GenCos may not always be able to produce the supply needed to satisfy the expected demand. Furthermore, in [Manvi *et al.*, 2024], an incremental case of prosumers placing multiple buy and sell bids is considered.

5.1 Adequate Supply Case

Let us define a few entities that will help to illustrate the proposed MPNE policy. First, denote $Q^{\mathcal{D}_{-b},h} = Q^{\mathcal{D},h} - Q^{b,h}$ as the demand of players excluding the player $b \in [N]$ at round h. Second, let u_h be the lowest index of an ask in the sorted set $\mathcal{L}^h$ such that the supply up to the first u_h asks is greater than the quantity $Q^{\mathcal{D},h} - Q_-^{\mathcal{S},h}$. That is, $u_h = \arg\min_i (Q^{\mathcal{D},h} - Q_-^{\mathcal{S},h} < \sum_{m=1}^{i} q_m^h)$. Similarly, let v_h^b be an index of the sorted set $\mathcal{L}^h$ such that the supply up to first v_h^b asks is greater than $Q^{-b,h} - Q_-^{\mathcal{S},h}$. That is,

$$v_h^b = \arg\min_j \left(Q^{\mathcal{D}_{-b},h} - Q_-^{\mathcal{S},h} < \sum_{m=1}^{j} \hat{q}_m^h \right) \;\forall\; b \in [N]. \tag{3}$$

Finally, define an index of $\mathcal{L}^h$ as $z_h = \max\{v_h^1, v_h^0\}$ and let ϕ^h be the player who bids at a price p_{z_h}, where v_h^0 is $u_h - (H-h)$ and ϕ^h is defined as,

$$\phi^h = \max\{1, \arg\max_b\{v_h^b \le v_h^0\}\}. \tag{4}$$

MPNE Policy for Adequate Supply Case: The joint policy π_* suggests that the player b needs to check $Q^{b,h} > q_{u_h}$, that is, whether her selling quantity is greater than the ask quantity at the index u_h (of $\mathcal{L}^h$). If yes, then the player places sell bid[4] at price $p_{-,*}^{b,h} = p_{u_h}$ with quantity $q_{-,*}^{b,h} = Q_-^{b,h}$. However, if $Q^{b,h} \le q_{u_h}$ the prosumer places sell bid at price $p_-^{b,h} = p_{u_h} - \epsilon$, where $\epsilon > 0$ is a small constant. Furthermore, the player has to check the condition $b = \phi^h$ and if the constraint is satisfied, then the player bids at buy bid price $p_{+,*}^{b,h} = p_{z_h}$ with quantity $q_{+,*}^{b,h} = Q_+^{b,h}$. However, when $b \ne \phi^h$, the player places a buy bid price $p_{+,*}^{b,h} = p_{\max}$ and quantity bid equal to her requirement. The MPNE policy $\pi_*^{b,h}$ with buy and sell bids $(p_{+,*}^{b,h},\ q_{+,*}^{b,h},\ p_{-,*}^{b,h},\ q_{-,*}^{b,h})$ is given in Equation (5). Note that the policy $\pi_*^{b,h}$ considers both sell bids and buy bids of the prosumer, unlike [Manvi and Subramanian, 2023], where only buy bids are considered.

$$\pi_*^{b,h} = \begin{cases} \left(p_{z_h},\ Q_+^{b,h}, p_{u_h} - \epsilon \mathbb{1}_{\{Q_-^{b,h} \le q_{u_h}\}},\ Q_-^{b,h}\right) & \text{if } b = \phi^h \\ \left(p_{\max},\ Q_+^{b,h}, p_{u_h} - \epsilon \mathbb{1}_{\{Q_-^{b,h} \le q_{u_h}\}},\ Q_-^{b,h}\right) & \text{o.w.} \end{cases} \tag{5}$$

Having explained the candidate policy π_*, the value function of the joint policy $\pi_*^{b,h}$ at round h for player $b \in [N]$ at state $s^h \in S$ is given in Equation (6) using the properties of the uniform clearing mechanism defined in Definition 1.

$$V^h_{\pi_*^b, \pi_*^{-b}}(s) = \begin{cases} \Big[p_{z_h} \cdot Q^h - p_{z_h} \cdot \sum_{i \in [N] \setminus \phi^h} q_j^{i,h} \\ \quad - p_{z_h} \cdot \beta_*^{b,h} + \sum_{k=h+1}^{H} p_{z_k} \cdot Q^k \\ \quad\quad - \sum_{k=h+1}^{H} p_{z_k} \cdot \beta_*^{b,h} \Big], \text{ if } b = \phi^h \\ p_{z_h} \cdot (Q^{buy,h} - \beta_*^{b,h}), \quad \text{o.w.} \end{cases} \tag{6}$$

Equilibrium Analysis: We demonstrate that the candidate policy in Equation (5) is indeed MPNE policy in the space of all deterministic policies. More specifically, we show that, for all players $b \in [N]$, for all state $s^h \in S$, for all round $h \in [H]$ and for any deterministic policy π^b, Equation (2) is satisfied. To this end, we consider the value function for all possible deviations where the sell bid deviations are tabulated in Table 1. Furthermore, the buy bid deviations are exactly the same as in Table 1 (except for the notation from $p_-^{b,h}$ to $p_+^{b,h}$). Note that, by assumption, prosumers cannot buy more than they require and sell more than they have; hence, only five sell and buy bid deviations are possible. Finally, the combined deviations $\pi^b \in \Pi^b$ is the cartesian product of the sell bid deviations and the buy bid deviations[5]. In the sequel, we provide a preliminary result in the Lemma 1. The first part of Lemma 1 provides a property of the clearing price λ^h for a uniform clearing rule, which satisfies Definition 1. The second part of Lemma 1 provides a condition on the cost of balancing outside the horizon of the PDA.

Lemma 1. *Given that the supply curve from bulk sellers across rounds $h \in [H]$ is constant and the clearing mechanism satisfies the properties in Definition 1, we have the following.*

1. *The clearing price at h satisfy $\lambda^h \ge \lambda^{h+1}$ for $h \in [H-1]$.*
2. *The condition on the balancing price to buy outside the auction at $H+1$ denoted by Υ is given by $\Upsilon > \gamma \cdot p_{\max}$, where $\gamma > 1$.*

[4] Bid indices omitted since there is only one sell and one buy bid.

[5] In total, $25(5 \times 5)$ deviations are possible.

Equal Priced Deviation	
$p_-^{b,h} = p_{-,*}^{b,h}, q_-^{b,h} < q_{-,*}^{b,h}$	
Higher Priced Deviations	Lower Priced Deviations
$p_-^{b,h} > p_{-,*}^{b,h}, q_-^{b,h} < q_{-,*}^{b,h}$	$p_-^{b,h} < p_{-,*}^{b,h}, q_-^{b,h} < q_{-,*}^{b,h}$
$p_-^{b,h} > p_{-,*}^{b,h}, q_-^{b,h} = q_{-,*}^{b,h}$	$p_-^{b,h} < p_{-,*}^{b,h}, q_-^{b,h} = q_{-,*}^{b,h}$

Table 1: Possible Sell Bid Deviations

Proof. Please refere to [Manvi *et al.*, 2024]. □

We now move to the Theorem 1, which shows that the value function of the deviated policies is greater than the candidate policy's value function.

Theorem 1. *Given the conditions in Lemma 1 hold, we have the following.*

1. *The sell bid deviations $\pi^b \in \Pi^b$ of player b with any buy bids $p_+^{b,h} \in [0; p_{\max}]$ and $q_+^{b,h} \in [0; Q_+^{b,h}]$ satisfy* (2).
2. *The buy bid deviations $\pi^b \in \Pi^b$ of player b with any sell bids $p_-^{b,h} \in [0; p_{\max}]$ and $q_-^{b,h} \in [0; Q_-^{b,h}]$ satisfy* (2).

Proof. Please refer to [Manvi *et al.*, 2024]. □

5.2 Inadequate Supply Case

We now consider the inadequate supply, (that is, $Q^{\mathcal{S},h} < Q^{\mathcal{D},h}$). To provide the MPNE for this case, we modify the definitions of u_h, ϕ^h, v_h^k and z_h provided in the adequate supply case. First, we assign $u_h = L^h$ as the value. Second, ϕ^h is given as

$$\phi^h = \arg\min_j \{Q^{\mathcal{S},h} \leq \sum\nolimits_{b=1}^{j} Q^{b,h}\} \tag{7}$$

Third, when $\phi^h = N$, we set index v_h^N as

$$v_h^N = \arg\min_j \left(Q^{\mathcal{D}_{-N},h} - Q_-^{\mathcal{S},h} \leq \sum_{m=1}^{j} q_m^h \right).$$

Finally, z_h when $\phi^h = N$ is defined as $z_h = \max\{v_h^{\phi}, v_h^0\}$. However, if $\phi^h < N$, then $p_{z_h} = p_{\max}$.

MPNE Policy for Inadequate Supply Case: The buy bids mentioned in Equation (5) now use modified definitions, whereas the sell bids change to bid at a price $p_-^{b,h} = \Upsilon - \epsilon$ and bid quantity $q_-^{b,h} = Q_-^{b,h}$. In [Manvi *et al.*, 2024], we provide a Lemma that shows the policy is indeed MPNE.

6 Experimental Evaluation

In the previous sections, we presented analytical equilibrium solutions for PDAs using a Markov game framework. In doing so, we had a complete information setup wherein the players (prosumers) have knowledge of the supply curve and the demand information of other players at each round of the PDA. In this section, we use the equilibrium solutions (5) obtained in the previous section to propose a bidding strategy MPNE-BBS for the incomplete information case where a buyer has neither the knowledge of the demand requirement of other players nor complete information about the supply curve. We start by providing the algorithm for MPNE-BBS, followed by numerical experiments to showcase the efficacy of the proposed strategy. The numerical experiments are conducted on the wholesale market module of the Power Trading Agent Competition (PowerTAC) [Ketter *et al.*, 2020].

PowerTAC: PowerTAC is an efficient and close-to real-world smart grid simulator that models all the crucial elements of a typical smart grid system, including GenCo and energy brokers acting as prosumers. During the simulation (or game), which typically lasts for 60 simulation days, an energy broker has to compete against several other brokers. To test our bidding strategy, we focus on PowerTAC's wholesale market PDAs, where the energy broker plays a crucial role in buying/selling energy. These PDAs are day-ahead auctions in which the broker can purchase energy 24 hours ahead of the delivery timeslot by participating in a total of 24 auctions at an interval of every hour. PowerTAC PDA employs ACPR and uniform pricing rules for clearing and payments. To bid in PowerTAC PDA, a broker must submit the bid price and quantity (decided based on demand forecasts). For determining the bid price, a broker may use the information available from the server, which includes the market-clearing price, its own cleared quantity, net cleared quantity and orderbook information. Orderbook includes an anonymized list of uncleared asks and bids, while the knowledge about the other brokers' cleared bids/asks is kept hidden. Failing to acquire the required quantity from the wholesale market, a broker has to purchase the remaining quantity at balancing market prices, which are typically high and serve as a penalty for a broker for causing imbalance. The simulation also includes a buyer called MISO that procures energy for a region that contains retailer users not serviced by the main brokers of the PowerTAC setup. The MISO buyer's energy requirement is almost ten times the PowerTAC retail market's energy requirements and substantially affects the clearing prices. The MISO buyer purchases all of its estimated demand in the first round of the PDA and any excess procurement is sold in the subsequent rounds. The MISO always places a market order to buy or sell energy in any round of the PDA. The GenCos are the primary sellers in the market that follow a quadratic cost function to decide the ask prices. Refer to PowerTAC specifications [Ketter *et al.*, 2020] for more information.

MPNE-BBS Algorithm: We propose an algorithm MPNE-BBS for the incomplete information setting of PowerTAC inspired by the equilibrium analysis on the complete information setting. As our proposed MPNE-BBS bidding algorithm uses some design ideas from VV21, we first describe the design of VV21. It models the cost supply curve of the GenCos from the uncleared ask information data available from the simulator. The idea is to locate the price corresponding to the broker's bid quantity (requires demand forecast of broker's and market's demands) on the supply curve, treat that price as the upper bound on the limit-prices, and place multiple bids below that price. The reason for placing multiple bids below the chosen upper bound is that it aims to procure the majority of the quantity from the asks of the MISO buyer and other prosumers in the market and treats GenCo as the sup-

Algorithm 1 MPNE-BBS

1: totalDmd[] $\leftarrow$ netDmdPredict(currentTime)
2: $Q^{b,h}$[] $\leftarrow$ indvDmdPredictor[currentTime]
3: **for** hour in $[1,\ldots,23]$ **do**
4: futureTime $\leftarrow$ currentTime + hour
5: unclearedAsks[] $\leftarrow$ Auction(currentTime-1,futureTime)
6: **if** unclearedAsks is not empty **then**
7: $Q^{\mathcal{S},h} \leftarrow$ sum(unclearedAsks.q)
8: $Q^{\mathcal{D},h} \leftarrow$ totalDmd[futureTime]
9: **if** $Q^{\mathcal{S},h} \geq Q^{\mathcal{D},h}$ **then**
10: $p_{u_h} \leftarrow \min p_r$ s.t $Q^{\mathcal{D},h} \leq \sum_{i=1}^{r}$ unclearedAsks.q_i
11: $Q^{\mathcal{D}_{-b},h} \leftarrow Q^{\mathcal{D},h} - Q^{b,h}$[futureTime]
12: $p_{v_h^b} \leftarrow \min p_r$ s.t $Q^{\mathcal{D}_{-b},h} \leq \sum_{i=1}^{r}$ unclearedAsks.q_i
13: **else**
14: $p_{u_h} \leftarrow p_{L^h}$, $p_{v_h^b} \leftarrow P_{L^h}$
15: **end if**
16: $v_h^0 \leftarrow \max\{1, u_h - \text{hour} + 1\}$
17: estBidPrice = $\max\{p_{v_h^0}, p_{v_h^b}\}$
18: **else**
19: clearedPrices[] $\leftarrow$ Auction(t,t+hour) $\forall t <$ currentTime
20: estBidPrice = max{cleredPrices}
21: **end if**
22: **if** (currentTime is far from futureTime) **then**
23: minP = α_f * estBidPrice; maxP = β_f * estBidPrice
24: **else**
25: minP = α_c * estBidPrice; maxP = β_c * estBidPrice
26: **end if**
27: Sample P prices, $p_i^{b,h} \sim$ U[minP,maxP]
28: Distribute $Q^{b,h}$ uniformly across P prices, $q_i^{b,h} \leftarrow Q^{b,h}/P$
29: bidList $\leftarrow (p_i^{b,h}, q_i^{b,h})\ \forall\, i \in \{1, 2, ..., P\}$
30: Auction(currentTime,futureTime) $\leftarrow$ bidList
31: **end for**

plier in the last resort. The supply curve of the GenCo also has an element of randomization between rounds of the PDA; hence, placing multiple bids in a range-bound manner helps the agent procure energy at a lower price. More details of the VV21 strategy can be found in [Chandlekar *et al.*, 2022a].

On the other hand, MPNE-BBS decides the price on the supply curve by using the forecasts of the total demand and the broker's demand to determine the indices u_h and v_h^b on the uncleared asks list (number of uncleared asks to satisfy the demand). Using these indices, it estimates the best possible limit-price, which follows the analytical solution presented in Section 5. After deciding the price, it places multiple bids similar to VV21 to work around the supply curve's randomness. Unlike VV21, MPNE-BBS aims to purchase most of the quantity from GenCo.

As shown in the algorithm, MPNE-BBS takes the current bidding timeslot as input, uses the 24th hour (the first opportunity) of each auction to observe the uncleared asks, and places a list of bids for each of the next 23 hours. For each of these future hours (futureTime), the algorithm queries for the list of uncleared asks from past auction data. These asks are a list of price and quantity tuples (p_i, q_i) sorted in increasing order of prices. Then, based on the list of unclear asks, the algorithm estimates the bid price (estBidPrice) by utilising the knowledge of the market and its own demand forecasts. The bid price estimation approach in lines 4 to 21 of algorithm 1 is inspired by the equilibrium solution presented in Equation (5). As the buyer has 24 opportunities to procure the required demand, thus can afford to take risks during the initial rounds and play conservatively in the last few rounds, as shown in lines 23 to 27. The hyperparameters α_f, β_f, α_c and β_c can be decided based on the risk level of the buyer. After sampling P prices uniformly, the buyer's required quantity is uniformly divided into P bids and submitted for clearing.

Benchmark: Below, we briefly describe the state-of-the-art as well as baseline bidding strategies of PowerTAC PDAs that are used in the experiments to compare the performance of MPNE-BBS. We use six bidding strategies to compete against MPNE-BBS, namely, VidyutVanika21 (VV21) [Chandlekar *et al.*, 2022a], Deep-Deterministic-Policy-Gradient-based Bidding Strategy (DDPG) [Chandlekar *et al.*, 2022b], SPOT [Chowdhury *et al.*, 2018], VidyutVanika18 (VV18) [Ghosh *et al.*, 2020], Zero Intelligence (ZI) and Zero Intelligence Plus (ZIP) [Cliff, 1997]. DDPG and VV18 are scale-based bidding strategies implemented using RL algorithms. SPOT is built on top of Monte Carlo Tree Search (MCTS). ZIP is a heuristic strategy that keeps a profit margin while bidding and the ZI is a randomized approach to placing bids by sampling prices from a uniform distribution. The reader is referred to Appendix 6 for more details about these strategies.

Experimental Setup: To test the efficacy of MPNE-BBS against the above-listed bidding strategies, we perform two sets of experiments on the PowerTAC platform, one that includes the MISO buyer and the other without it. In these experiments, we compare the unit purchase costs of the brokers in the wholesale market, a strategy having a lower purchase cost is preferable. In Set-1, we play all-player games that include the above six bidding strategies and MPNE-BBS, along with MISO buyer in the market. We ask all the players except MISO to procure a fixed demand requirement for each timeslot; the demand requirement is the same for all the brokers for a fair comparison. Set-1 is further divided into four configurations depending on the demand requirements of the brokers, namely, low-demand, mid-demand, high-demand and extreme-demand. Similarly, in Set-2, we remove the MISO buyer from the market and repeat the same experiments for all four demand levels. We play 10 games for each configuration in each set (Figures 1 and 2). Additionally, we play two-player games between MPNE-BBS and all the other strategies; that is, we play 10 games between MPNE-BBS and VV21, 10 games between MPNE-BBS and ZI, and so on. We perform these experiments for both the sets and all four demand levels as mentioned above (Tables 2 and 3).

Results: As shown in Figure 1, MPNE-BBS is one of the best-performing bidding strategies in terms of wholesale cost as it achieved close to the best or second-best cost among the seven brokers across variable demand levels. Specifically, its wholesale cost is close to the best wholesale cost in the market for extreme demand level (only 6% away from the best) and close to second-best for high and mid demand levels (4.7% and 3.4% difference, respectively), with only VV21 performing better. The performance is more prominent in Set-2; as shown in Figure 2, it achieves a wholesale cost very

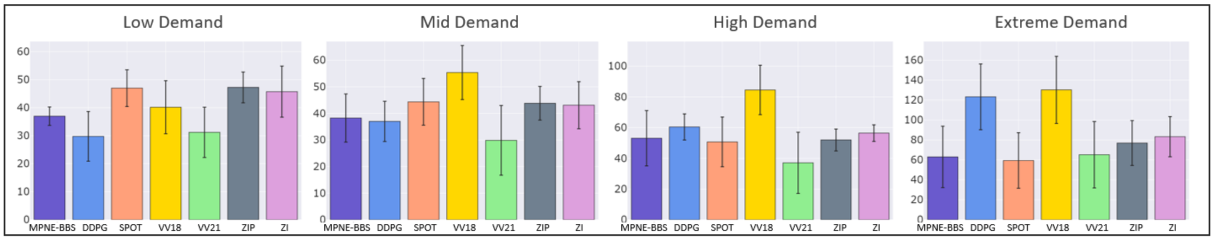

Figure 1: Wholesale Cost Comparison in 7-Player Games (with MISO Buyer)

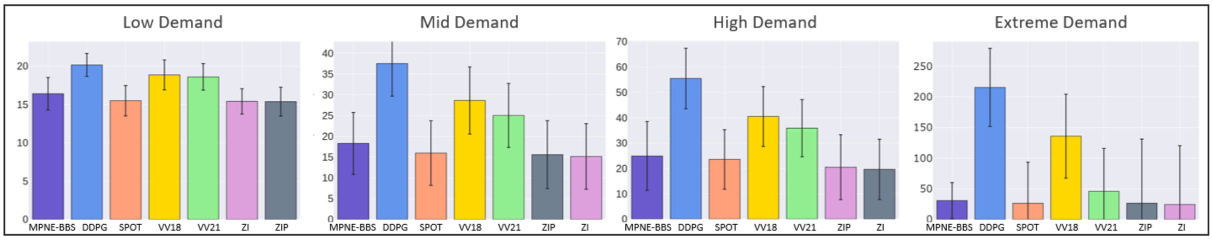

Figure 2: Wholesale Cost Comparison in 7-Player Games (without MISO Buyer)

Opponent	**Low**	**Mid**	**High**	**Extreme**
VV21	**1.01**	**0.98**	**1.03**	**1.18**
DDPG	0.73	0.85	**1.12**	**1.64**
SPOT	**1.45**	**1.09**	0.87	0.77
VV18	**1.32**	**1.36**	**1.49**	**1.76**

Table 2: Performance in 2-Player Games (with MISO Buyer)

Opponent	**Low**	**Mid**	**High**	**Extreme**
VV21	**0.99**	**0.99**	**1.0**	**0.99**
DDPG	**1.08**	**1.07**	**1.02**	**1.08**
SPOT	0.91	0.82	**1.08**	**1.09**
VV18	**1.04**	**1.15**	**1.16**	**1.13**

Table 3: Performance in 2-Player Games (without MISO Buyer)

close to the best wholesale cost in the market across all demand levels and consistently outperforms DDPG and VV21.

The results of the 2-Player experiments in Tables 2 and 3 show the wholesale cost of the opponent strategy relative to MPNE-BBS, a value more than 1 would indicate MPNE-BBS is the superior bidding strategy among the two. Particularly, it matches VV21, which is the best strategy in the literature and achieves similar wholesale cost as VV21 in all demand levels, with and without MISO buyer. It even outperforms VV21 for the extreme demand level where VV21's wholesale cost is 1.18 times higher than MPNE-BBS's cost. Overall, the results show that it achieves superior performance against each opponent and for almost all demand levels. This set of experiments aims to validate the efficacy of MPNE-BBS against several different state-of-the-art strategies having different bidding patterns. Furthermore, we performed the same experiments for 3-Player and 5-Player games. On an average, the relative costs of best and worst opponent strategies were 1.13 and 1.87, respectively in 5-player without MISO games; 1.0 and 2.34, respectively in 5-player with MISO games. Similarly, for 3-player games, these numbers are 1.27 and 1.47, respectively for best and worst opponent strategies without MISO; 0.98 and 1.19, respectively with MISO. Thus, the simulation results show that the MPNE-BBS achieves significant performance improvements against the best state-of-the-art bidding strategy across various number of and players in a game for various market and demand scenarios.

7 Conclusion

In this paper, we proposed equilibrium strategies for prosumers involved in a PDA to buy and sell commodities. We modelled the PDA as a Markov game with prosumers as players and derived equilibrium solutions for the complete information setting when the players are aware of the supply curve and the outstanding demand requirement of other players at every round of the PDA. Specifically, we derived MPNE solutions for the setting when the players compete to procure required commodities. Thereafter, the proposed MPNE solutions were used to design a bidding algorithm called MPNE-BBS for a more practical setup and its efficacy was demonstrated using the PowerTAC simulator test-bed against several state-of-the-art algorithms. In future work, we hope to extend the analysis of equilibrium solutions to the incomplete information setting by modelling the PDA as a partially observable stochastic (Markov) game and compare its effectiveness against the algorithm proposed in this work.

References

[Angaphiwatchawal *et al.*, 2021] Pikkanate Angaphiwatchawal, Chuppawit Sompoh, and Surachai Chaitusaney. A Multi-k Double Auction Pricing Mechanism for Peer-to-Peer Energy Trading Market of Prosumers. In *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE, May 2021.

[Chandlekar *et al.*, 2022a] Sanjay Chandlekar, Bala Suraj Pedasingu, Easwar Subramanian, Sanjay Bhat, Praveen Paruchuri, and Sujit Gujar. VidyutVanika21: An Autonomous Intelligent Broker for Smart-grids. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 158–164, 7 2022.

[Chandlekar *et al.*, 2022b] Sanjay Chandlekar, Easwar Subramanian, Sanjay Bhat, Praveen Paruchuri, and Sujit Gujar. Multi-Unit Double Auctions: Equilibrium Analysis and Bidding Strategy Using DDPG in Smart-Grids. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '22, page 1569–1571, 2022.

[Chatterjee and Samuelson, 1983] Kalyan Chatterjee and William Samuelson. Bargaining Under Incomplete Information. In *Operations Research*, volume 31, 1983.

[Chowdhury *et al.*, 2018] Moinul Morshed Porag Chowdhury, Christopher Kiekintveld, Son Tran, and William Yeoh. Bidding in Periodic Double Auctions Using Heuristics and Dynamic Monte Carlo Tree Search. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 166–172, 7 2018.

[Cliff, 1997] Dave Cliff. *Minimal-Intelligence Agents for Bargaining Behaviors in Market-Based Environments*. Technical Report HPL-97-91, Hewlett Packard Labs., 1997.

[Constantinides and Cartlidge, 2021] Theodoros Constantinides and John Cartlidge. Block auction: A general blockchain protocol for privacy-preserving and verifiable periodic double auctions. In *2021 IEEE International Conference on Blockchain (Blockchain)*, pages 513–520, 2021.

[Disbrey and Zeier, 2020] Stuart Disbrey and Irene Zeier. Nord Pool AS Anual Report, 2020. [Online; accessed 16-January-2024].

[Du *et al.*, 2021] Yan Du, Fangxing Li, Helia Zandi, and Yaosuo Xue. Approximating Nash Equilibrium in Day-ahead Electricity Market Bidding with Multi-agent Deep Reinforcement Learning. *Journal of Modern Power Systems and Clean Energy*, 9(3):534–544, 2021.

[Ghasemi *et al.*, 2020] Arman Ghasemi, Amin Shojaeighadikolaei, Kailani Jones, Morteza Hashemi, Alexandru G. Bardas, and Reza Ahmadi. A Multi-Agent Deep Reinforcement Learning Approach for a Distributed Energy Marketplace in Smart Grids. In *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, November 2020.

[Ghosh *et al.*, 2020] Susobhan Ghosh, Sujit Gujar, Praveen Paruchuri, Easwar Subramanian, and Sanjay Bhat. Bidding in Smart Grid PDAs: Theory, Analysis and Strategy. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(02):1974–1981, Apr. 2020.

[Ketter *et al.*, 2020] Wolfgang Ketter, John Collins, and Mathijs de Weerdt. The 2020 Power Trading Agent Competition. *SSRN Electronic Journal*, 2020.

[Krausz and Rieder, 1997] Alexander Krausz and Ulrich Rieder. Markov Games with Incomplete Information. *Mathematical Methods of Operations Research*, 46:263–279, 1997.

[Kuate *et al.*, 2013] Rodrigue Kuate, Minghua He, and Maria Chli. An Intelligent Broker Agent for Energy Trading: An MDP Approach. pages 234–240, 08 2013.

[Manvi and Subramanian, 2023] Bharat Manvi and Easwar Subramanian. A Nash Equilibrium Solution for Periodic Double Auctions. In *2023 IEEE 62nd Conference on Decision and Control (CDC)*, 2023. Preprint available at Arxiv.

[Manvi *et al.*, 2024] Bharat Manvi, Sanjay Chandlekar, and Easwar Subramanian. Optimizing Prosumer Policies in Periodic Double Auctions Inspired by Equilibrium Analysis (Extended Version), 2024.

[Orfanoudakis *et al.*, 2021] Stavros Orfanoudakis, Stefanos Kontos, Charilaos Akasiadis, and Georgios Chalkiadakis. *Aiming for Half Gets You to the Top: Winning PowerTAC 2020*, pages 144–159. 07 2021.

[Parsons *et al.*, 2011] Simon Parsons, Juan A. Rodriguez-Aguilar, and Mark Klein. Auctions and Bidding : A Guide for Computer Scientists. *ACM Computing Surveys*, 43(2):1–59, January 2011.

[Rashedi *et al.*, 2016] Navid Rashedi, Mohammad Amin Tajeddini, and Hamed Kebriaei. Markov Game Approach for Multi-agent Competitive Bidding Strategies in Electricity Market. *IET Generation, Transmission and Distribution*, 10(15):3756–3763, November 2016.

[Satterthwaite and Williams, 1989] Mark A. Satterthwaite and Steven R. Williams. Bilateral Trade with the Sealed Bid k-double Auction: Existence and Efficiency. *Journal of Economic Theory*, 48(1):107–133, June 1989.

[Taylor, 2015] Joshua Adam Taylor. *Convex Optimization of Power Systems*. Cambridge University Press, 2015.

[Tesauro and Bredin, 2002] Gerald Tesauro and Jonathan L. Bredin. Strategic Sequential Bidding in Auctions Using Dynamic Programming. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 2*, AAMAS '02, page 591–598, 2002.

[Urieli and Stone, 2014] Daniel Urieli and Peter Stone. Tactex'13: A champion adaptive power trading agent. In *Association for the Advancement of Artificial Intelligence (AAAI)*, 2014.

[Vetsikas, 2014] Ioannis A. Vetsikas. Equilibrium Strategies for Multi-Unit Sealed-Bid Auctions with Multi-Unit Demand Bidders. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*, AAMAS '14, page 1053–1060, 2014.

[Wikipedia contributors, 2023] Wikipedia contributors. Auction – Wikipedia, the free encyclopedia, 2023. [Online; accessed 5-December-2023].

[Wilson, 1992] Robert B. Wilson. Chapter 8 Strategic Analysis of Auctions. *Handbook of Game Theory With Economic Applications*, 1:227–279, 1992.

[Yang and Wang, 2020] Yaodong Yang and Jun Wang. An Overview of Multi-Agent Reinforcement Learning from Game Theoretical Perspective. *CoRR*, abs/2011.00583, 2020.

[Zhang *et al.*, 2023] Yizhou Zhang, Guannan Qu, Pan Xu, Yiheng Lin, Zaiwei Chen, and Adam Wierman. Global Convergence of Localized Policy Iteration in Networked Multi-Agent Reinforcement Learning. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 7(1):1–51, February 2023.