Decentralized Transactive Energy Auctions with Bandit Learning

Zibo Zhao
School of Industrial Engineering
Purdue University, West Lafayette, Indiana, USA

Kiyoshi Nakayama, Ratnesh Sharma
Energy Management Department
NEC Laboratories America, California, USA

***Abstract*—The power systems worldwide have been embracing the rapid growth of distributed energy resources. Commonly, distributed energy resources exist in the distribution level, such as electric vehicles, rooftop photovoltaic panels, and home battery systems, which cannot be controlled by a centralized entity like a utility. However, a large number of distributed energy resources have potential to reshape the power generation landscape when the owners (prosumers) are allowed to send electricity back to the grids. Transactive energy paradigms are emerging for orchestrating the coordination of prosumers and consumers by enabling the exchange of energy among them. In this paper, we propose a transactive energy auction framework based on blockchain technology for creating trustworthy and transparent transactive environments in distribution networks, which does not rely on a centralized entity to clear transactions. Moreover, we propose intelligent decentralized decision-making strategies by bandit learning for market participants to locally decide their energy prices in auctions. The bandit learning approach can provide market participants with more benefits under the blockchain framework than trading energy with the centralized entity, which is further supported by the preliminary simulated results conducted over our blockchain-based platform.**

***Index Terms*—transactive energy auctions, distributed energy resources, decentralized decision-making, bandit learning, blockchain**

I. Introduction

The current power distribution systems are under an overwhelming trend of rapidly growing of distributed energy resources (DERs), such as rooftop photovoltaic panels, electric vehicles, and stationary battery storages. The low-voltage energy provided by DERs is more flexible and can be remotely controllable as part of the Internet of Things. Therefore, DERs are expected to be adopted intelligently to reduce costs, offset volatility, and integrate more renewable resources through a coordination mechanism among autonomous prosumers [1]. Transactive energy (TE) provides a promising solution for effectively incentivizing DERs [2], [3], [4], in which DER operational decisions are made based on market value information. However, in the current distribution systems, prosumers can only sell energy to the centralized distribution system operator (DSO) like a utility who could be a monopoly for organizing an inefficient market. The small-scale prosumers are disadvantaged in the negotiation with DSO regarding the payments for their DER services, which may discourage the adoption of DERs [5].

In this work, we propose a blockchain-based TE market paradigm which organizes an efficient and trustworthy trading market with the assistance of asymmetric encryption, digital signature, distributed and consensus mechanism [6]. With smart contracts, transactions among untrustworthy entities can be executed automatically under predefined conditions, which is independent of any third entity. There has been limited work about deployment of blockchain in energy transaction field [7], [8], [9], [10], [11]. Only [10], [11] study into utilizing smart contract to perform a distributed auction for energy market. The work in [10] considers independent Vickrey auctions whenever a seller has some energy to offer, which makes trading activities intractable for buyers, and no intelligent bidding strategies is given. An adaptive quotation strategy for participants is proposed in [11] which is heuristic and does not consider opponents' animosity when all participants attend the same auction. In this study, we introduce intelligent pricing strategies by bandit learning over transaction history, by which distributed agents can locally decide its quotation price without knowing others' information and achieve higher expected rewards in auctions than trading with centralized utility. The animosity of opponents can be taken care by the interaction dynamics of the multi-armed bandit (MAB) game framework [12] when every agent is solving its MAB problem.

II. Double Auctions over Blockchain-based Energy Trading Platform

A. Blockchain-based Platform Infrastructure

The proposed blockchain-based TE market platform in a distribution system consists of three entities including prosumers, common consumers, and DSO. Since deposit and bill statements will be settled according to metering values at the end of each transaction window, we assume that participants have smart meters communicating their usage and generation readings with smart contracts. Every participant maintains an account with unique encrypted address for signed transactions that are settled through crypto-currencies (e.g. 1 Ether $\approx$ \$115 at the time of writing).

Figure 1 describes the trading platform for the transactive energy framework utilizing blockchain. As shown in Figure 1, smart contracts are converted into binary data and stored in the chain. Using contract modifiers, smart contracts can describe who can access which functions and what functions can modify which data. As in Figure 1, the smart contract mainly has five functions: Quotes Submission, Auction Clear, Payment Delivery, Smart Metering, and Deposit &

Bill Settlement. The detail of each function is explained in Section IV. Access to each function is restricted depending on function callers. For instance, Quotes Submission can only be accessed by prosumers and consumers whereas Deposit & Bill Settlement is only conducted by utilities as the function is about utility settlement process. In particular, utilities can contribute to Auction Clear only if there are not sufficient votes from the participants or the votes are equally dispersed. Also, each function is clearly mapped to its data storage in the smart contract while time index and address data are utilized whenever they are needed by the functions listed in Figure 1.

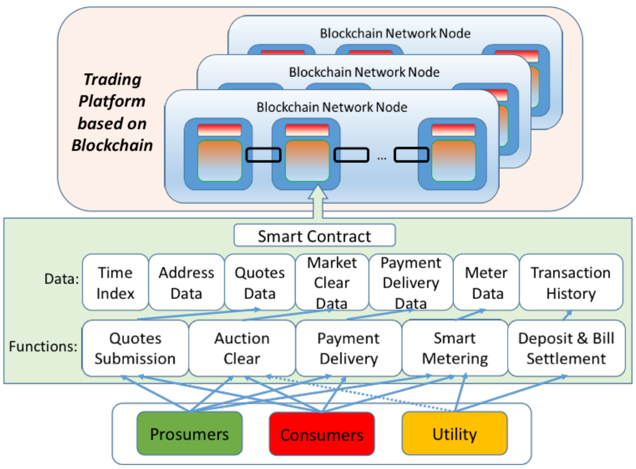

Figure 1. Blockchain-based TE Market Platform

B. Double Auction Mechanism

The transactive energy infrastructure involves interactions between sellers and buyers by holding a forward local energy market. The competitive equilibrium model by double auction is able to clear a market at the equilibrium price and achieve a balance between supplies and demands [13], [14]. When energy sellers and buyers interact their bids and offers, the equilibrium (Q^*, P^*) will be at the intersection of the supply and demand curves. As shown in Figure 2, the supply curve is formed by sellers' quotes in ascending order of their offering prices and the demand curve is formed by buyers' quotes in descending order of their bidding prices. The consumer agents whose bid is not lower than the equilibrium price P^* are trading at the equilibrium price. Similarly, all the supplier agents whose offer is not higher than the equilibrium are trading at the equilibrium price. The time-of-use (TOU) and feed-in-tariff (FIT) rates, which will be used in the utility settlement phase later, would be the baseline cases for buyers and sellers, respectively. Therefore, TOU and FIT become the practical boundaries for quotes in the auction as in Figure 2.

The blockchain-based distributed mechanism enables any authorized node in the network to clear the market without depending on any third authority as an auctioneer, and thus the equilibrium matching double auction mechanism can achieve an efficient market with the optimal social welfare given the submitted bids and offers [15], [16].

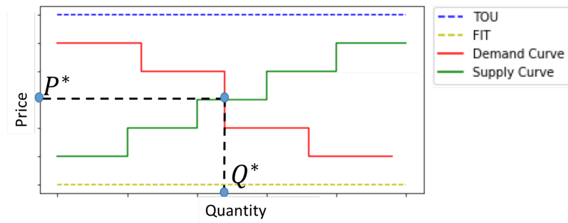

Figure 2. Market clear by equilibrium matching double auction for a transaction window

III. Decentralized Applications for Transactive Energy Auctions

A. Energy Quantity Forecast

To decide the energy quantity to bid or offer in an auction, participants conduct load and generation forecasting. Since auctions are held for the near future, we assume that participants can accurately forecast their energy amount to bid or offer by using their historical data from the same time slot of the previous days with similar weather conditions. The forecasting model is beyond the scope of this work. Herein, we simply let $f_i^{t,w}(\mathbf{X})$ denote the energy quantity forecast model of a participant i in time slot t under weather condition w, where $\mathbf{X}$ is the set of model inputs. Prior to the auction of transaction window t in day d, the participant i forecasts its energy quantity $\widehat{q}_i^{d,t} \sim f_i^{t,w}(\mathbf{X})$ by using the historical data, where positive values are for selling and negative values are for buying. As mentioned, the forecast is conducted for the near future, the expected forecast error $e_i^{d,t}$ and associated risk can be small for participants. Once the participant i submits the quote with $\widehat{q}_i^{d,t}$, he has to commit to the amount. The deviated amount by forecast error has to be exercised with the DSO with the high voltage system in the wholesale market. Thus, we can have the following balance equation for the distribution system in real time:

$$\textstyle\sum_{i \in \mathcal{B} \cup \mathcal{S}} (\widehat{q}_i^{d,t} + e_i^{d,t}) - Q_{Grid}^{d,t} = 0 \tag{1}$$

where $\mathcal{B}$ and $\mathcal{S}$ denote the buyers set and sellers set, respectively, and $Q_{Grid}^{d,t}$ is the energy injected into (when positive) or withdrew from (when negative) a bulk power grid by the DSO in real time.

B. Pricing Strategy with Bandit Learning

Without the transactive auctions, the agents can only buy at TOU P_u from the DSO or sell to the DSO at FIT P_f. The auctions under transactive energy paradigms enable market participants to exchange energy instead of trading with the DSO at FIT or TOU rate, which could improve participants' benefits with more competing rates, i.e. higher rates than FIT for sellers (here we consider zero marginal generation cost for DERs) and lower rates than TOU for buyers. Therefore, any rate between FIT and TOU is attracting to both buyers and sellers, and thus any reasonable agent will select its bidding or offering price $p_i^t \in (P_f^t, P_u^t)$ to compete against other agents in the auction of transaction window t, as shown in Figure 2.

This type of problem is considered as a MAB problem as it is about how to pick up the best price option to improve expected cumulative rewards across the transactions. Refer to [17], [18] for the overview of MAB problem. Consider days $(1,\ldots,D)$ and in each day there are transaction windows $(1,...,T)$. Then, the expected cumulative reward $E[R_i^{\Sigma}]$ for a participant i is as below in (3):

$$R_i^{d,t} = \underbrace{p_{clr}^{d,t} q_{i,clr}^{d,t}}_{\text{reward from auction}} + \underbrace{p_{i,unclr}^{d,t} q_{i,unclr}^{d,t}}_{\text{reward from DSO}} \tag{2}$$

$$E[R_i^{\Sigma}] = E[\textstyle\sum_{d=1}^{D} \sum_{t=1}^{T} R_i^{d,t}] \tag{3}$$

where $p_{clr}^{d,t}$ is an auction clear price (P^*) of the transaction window t in day d which is decided by all participants' submitted quotes $(\widehat{q}_i^{d,t}, p_i^{d,t})$. The quantities $q_{i,clr}^{d,t}$ and $q_{i,unclr}^{d,t}$ respectively are the accepted and unaccepted energy amount of $\widehat{q}_i^{d,t}$ in the participant i's quote cleared by the auction. Then, we can have:

$$q_{i,clr}^{d,t} + q_{i,unclr}^{d,t} = \widehat{q}_i^{d,t} \tag{4}$$

If $q_{i,unclr}^{d,t} > 0$, the energy is sold to a utility at FIT, i.e. $p_{i,unclr}^{d,t} = P_f^t$. Otherwise, the energy is purchased from a utility at TOU, i.e. $p_{i,unclr}^{d,t} = P_u^t$. As shown in Equation 2, the reward received in each transaction window consists of the revenues (or costs) from the auction and the DSO.

Therefore, each participant's reward is dependent on the auction clear result which is decided by all participants' quotes. A recent breakthrough in [12] about MAB-game provides good theoretical foundations to deal with the interactions among the participants' quotes. The work in [12] has proved that when all agents in a system are solving its own MAB problem using regret-minimization strategies, the interactions will lead to a steady state with a stationary population profile. In this context, when participants make decentralized decisions by solving interlinked MAB problems to price their energy against others' quotes, their expected rewards can converge to a level higher than trading with the DSO, which is shown in our preliminary simulation results in Section V. Classic bandit learning algorithms such as *UCB1*, *EXP3*, and their variants [17], [18] can be applied to pricing strategies of each participant by considering price options as action arms in the auction game. Through bandit learning, participants will have their expected rewards converged. We describe how a participant applies *UCB1* and *EXP3* for pricing its energy in the auctions of MAB games, respectively, in Algorithm 1 and Algorithm 2. In Algorithm 1 of *UCB1*, price options in each transaction window have their own history that consists of two elements, the number of time of being chosen and empirical mean reward. In Algorithm 2 of EXP3, a list of weights for the K price options is maintained for each transaction window and is used to compute probability distribution of price options.

Since each transaction window is independent from other windows in a day, each transaction window across multiple days is considered as a separate series of auction games. The participant applies the algorithms to learning over the price options of each transaction window according to previous days' auction clear results. The reward received from each auction is calculated as in Equation 2 and normalized by the best and worst case as shown in Equation 5. For a seller in transaction window t, the best case is that its quote is fully accepted in the auction and exercised at the highest price option $\overline{P} = max\{p \in \mathbb{Z} | P_f^t < p < P_u^t\}$. In contrast, for a buyer, the best case is fully buying energy at the lowest price option $\underline{P} = min\{p \in \mathbb{Z} | P_f^t < p < P_u^t\}$. The worst case for sellers and buyers is that the quote is not accepted and the energy is transacted in the normal utility settlement phase.

Algorithm 1 UCB1

Parameter: real $\sigma \in [0,1]$
Initialization: $\bar{r}_j^t = 0$ for $j = 1,\ldots,K_t$ and $t = 1,\ldots,T$.

Set $d \leftarrow 1$
while True **do**
  for $t = 1,\ldots,T$ **do**
    if $d \le K_t$ **then**
      Randomly draw a random price option j^* that has not been drew before from K_t price options.
      Receiving initial value for $\bar{r}_{j^*}^t \in [0,1]$.
      Set $n_{j^*}^t \leftarrow 1$.
    else
      Draw price option $j^* := \arg\max_j \left(\bar{r}_j^t + \sqrt{(2\sigma \ln d)/n_j^t}\right)$ from K_t options.
      Receiving normalized reward $r_{j^*}^t(d) \in [0,1]$.
      Update $\bar{r}_{j^*}^t$ for selected price option j^*.
      Set $n_{j^*}^t \leftarrow n_{j^*}^t + 1$.
    end if
  end for
  $d \leftarrow d+1$
end while

Algorithm 2 EXP3

Parameter: real $\gamma \in (0,1]$
Initialization: $w_j^t = 1$ for $j = 1,\ldots,K_t$ and $t = 1,\ldots,T$.

Set $d \leftarrow 1$
while True **do**
  for $t = 1,\ldots,T$ **do**
    Set $p_j \leftarrow (1-\gamma) w_j^t / \sum_{k=1}^{K_t} w_k^t + \gamma / K_t$ for $j = 1,\ldots,K_t$.
    Draw price option j^* randomly by the distribution of p_j.
    Receiving normalized reward $r_{j^*}^t(d) \in [0,1]$.
    Set $w_{j^*}^t \leftarrow w_{j^*}^t \exp\left(\gamma r_{j^*}^t(d) / (K_t p_{j^*})\right)$
  end for
  $d \leftarrow d+1$
end while

$$r^t(d) = \begin{cases} (R_i^{d,t} - \widehat{q}_i^{d,t} P_f^t)/(\widehat{q}_i^{d,t}(\overline{P} - P_f^t)) & \text{if } i \text{ is a seller} \\ (R_i^{d,t} - \widehat{q}_i^{d,t} P_u^t)/(\widehat{q}_i^{d,t}(\underline{P} - P_u^t)) & \text{if } i \text{ is a buyer} \end{cases} \tag{5}$$

IV. Transaction Settlement with Blockchain

The blockchain-based transactive energy architecture implements five essential steps to conduct settlement for each transaction window, including (1) quotes submission, (2) auction clear, (3) payment delivery, (4) smart metering, and (5) deposit and bill settlement. The specific process of each step is described below:

1) Quotes Submission: Prior to an auction, participants make decentralized decisions for their quote $(\widehat{q}_i^{d,t}, p_i^{d,t})$ as described in Section III-A and Section III-B. Then, along with submitting quotes, participants need to transfer deposit from their account to smart contracts. The deposit amount $x_i^{d,t}$ insures the quantity $\widehat{q}_i^{d,t}$ at TOU, i.e. $x_i^{d,t} = P_u^t |\widehat{q}_i^{d,t}| > p_{clr}^{d,t} |\widehat{q}_i^{d,t}|$.

The deposits are held by smart contact for enforcing commitment by participants, and thus can prevent fraudulent quotes. The quotes and associated deposits should be submitted within the predefined period and sealed with an encrypted address for privacy protection.

2) Auction Clear: The auction clear period starts right after the quote submission ends. As the quotes are revealed and advertised to the blockchain network, all the participants in the network can clear and verify the double auction on their own. In smart contract, we have a storage space for cleared price and quantity pair information at each time window of the auction. Authorized market participants perform auction clear as described in Section II-B through the predefined smart contract and publish their results within a time period. Market results of different nodes cleared by the smart contract shall be identical for the same set of quotes with the consistent auction clearing mechanism. However, if someone in the market believes it is not accurate or fair, that participant can post another cleared market price and quantity info pair in the smart contract. If one of the pairs of cleared market info receives more than 50% of votes, at that moment, that pair of price and quantity becomes the final cleared market information. Otherwise, the pair that receives the most votes becomes the cleared market info pair.

3) Payment Delivery: Once the price $p_{clr}^{d,t}$ and quantities $(q_{i,clr}^{d,t}, q_{i,unclr}^{d,t})$ are decided, the smart contracts deployed by authorized nodes match up quotes with peer-to-peer payment delivery. The accepted quotes of buyers and sellers are sorted in decreasing order of their absolute transacted monetary value $a_i^{d,t} = |p_{clr}^{d,t} q_{i,clr}^{d,t}|$, respectively. Then, the first buyer b transfers the amount of $max(a_b^{d,t}, a_s^{d,t})$ to the first seller s. Once a buyer pays up or a seller is paid up, it is removed from the queue. The total payment sent by buyers is the same as the total payment received by sellers. For buyers, if payments are transferred without default, the deposit will be returned. Otherwise, the default amount will be charged from the deposit and the rest will be returned. The sellers' deposit will still be held at this step. The monetary transactions can be separated from the OPF dispatch that is conducted by the DSO. It can also be solved by Smart Contract with blockchain, which is a potential future research direction.

4) Smart Metering: In real time, the sellers inject energy into the system. The DSO is responsible for dispatching the mismatch of supplies and demands based on the values of smart meters recorded in the smart contract in the blockchain network. If a seller is generating less than the committed quantity, the deviated amount will be charged at TOU by the DSO from its deposit and the utility will make up the deviation for demand-supply balance. Thus, the buyers' benefit will be insured by the sellers' deposit. Any generation surplus or demand deficit will be exercised at FIT with the DSO.

5) Deposit and Bill Settlement: At the end of the transaction window, the readings from smart meters are compared with the auction clear result, and the corresponding bill statements are generated for market participants. Any quantity deviation from the submitted quotes will be settled at TOU or FIT with the DSO. The deposit will be returned after accounting for default.

V. Numerical Results

We implement the proposed TE market and quotation strategies on the simulated network of the 56-bus sample test feeder [19] as shown in Figure 3. Bus 1 is used as the reference bus, and represents the DSO in the network. There are 100 prosumers and 100 consumers who are randomly distributed over the network.

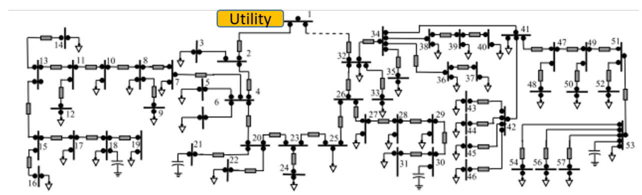

Figure 3. The 56-bus sample test feeder

We consider the simulation for the same season of 100 days for 3 years, and each day consists of 24 hourly transaction windows. Therefore, there are 24 series of 300-round auction games. Each participant randomly chooses from *UCB1-tuned, UCB1-normal, UCB2, ϵ-Greedy, EXP3* [17], [18] for selecting a bid or offer price. The described smart contracts are deployed in Ethereum nodes in a private blockchain network. Figure 4 shows the average normalized reward of transactions in last 30 days for each buyer and seller. After 300 days of learning, all sellers have similar rewards about 65%, and buyers' rewards are all around 40%.

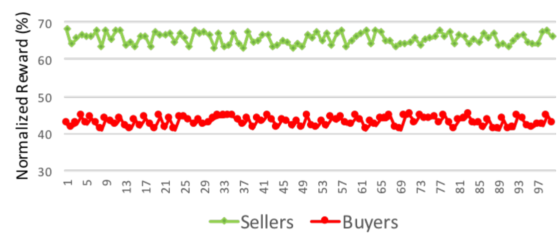

Figure 4. Participants' average normalized reward (%) of last 30 days

To more specifically illustrate the results above, we take an example of the peak hour 17:00-18:00 across 300 days. The TOU and FIT of the hour is ¢15/kwh and ¢9/kwh, respectively, and participants choose a price from $\{10, 11, 12, 13, 14\}$(¢/kwh) for each auction. All sellers individually forecast their offering quantity for each auction of the hour by sampling from the Beta distribution $30+20Beta(2,2)$ kwh with 5% forecasting error. Similarly, all buyers individually sample their bidding amount from the uniform distribution $Uniform(40, 60)$ kwh with 5% forecasting error. The total clear quantity Q^* and the total bidding and offering amounts of 300 auctions are shown in Figure 5(a) and 5(b), respectively. We can find that the total demand is beyond the total supply in the TE market, and the auction clear level converges to the supply level, which indicates that the energy supplied by distributed generators to the market is fully utilized with learning strategies for quotations.

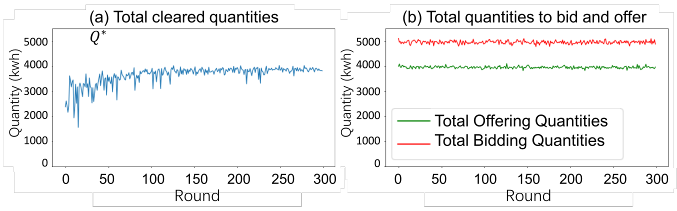

Figure 5. The total quantity (kwh) results of the auctions for hour 17:00-18:00: (a) Total clear energy quantities; (b) Total offering and bidding energy quantities.

Further, we present the results of Buyer 51 and Seller 51 using *EXP3*($\gamma = 0.2$) and *UCB1*($\sigma = 0.5$), respectively, in Figure 6. All the results ((b) - (d)) are shown as a moving average of 30 auctions. In (b), we can find that due to the supply shortage, in the later phase of learning, Buyer 51's cleared bidding energy converged around 80% on average whereas Seller 51's cleared offering quantity converged around 100%. Therefore, the expected normalized reward of Buyer 51 (i.e. 0.3) is lower than the Seller 51's normalized reward (i.e. 0.8) as shown in (c). However, Buyer 51 still achieves 10% expected improvement (i.e. cost decrease) over buying at TOU while Seller 51 achieves beyond 40% expected improvement (i.e. revenue increase) over selling at FIT as in (d). Hence, we can see through the bandit learning process in auction games that participants can learn about market conditions efficiently and have their expected rewards converged.

Figure 6. The simulated results of Buyer 51 and Seller 51 in the auctions: (a) Quoted quantities (kwh); (b) Clear percentages (%); (c) Normalized rewards (%); (d) Improvements over trading with the DSO (%).

VI. Conclusion and Future Work

A blockchain-based market design is proposed in this work for supporting exchange of energy among market participants in TE systems. Smart contracts that contain the double auction mechanism adopted by distributed nodes enable market clearing to be independent of any third authority, which leads to an efficient and trustworthy market for all participants. The proposed intelligent pricing strategies by bandit learning can provide participants with more benefits in auctions through learning over transaction history.

Distribution grid constraints are not considered currently. In the future, with more DERs injecting power into network, physical constraints like congestion shall be considered in deciding power delivery. In addition, smart contracts need to be designed to cope with optimal power flow problems by distributed nodes efficiently in a TE system.

References

[1] Y. Zhou, S. Ci, N. Lin, H. Li, and Y. Yang, "Distributed energy management of p2p energy sharing in energy internet based on cloud energy storage," in *Proceedings of the Ninth International Conference on Future Energy Systems*, pp. 173–177, ACM, 2018.

[2] K. Kok and S. Widergren, "A society of devices: Integrating intelligent distributed resources with transactive energy," *IEEE Power and Energy Magazine*, vol. 14, no. 3, pp. 34–45, 2016.

[3] A. Laszka, A. Dubey, M. Walker, and D. Schmidt, "Providing privacy, safety, and security in iot-based transactive energy systems using distributed ledgers," in *Proceedings of the Seventh International Conference on the Internet of Things*, p. 13, ACM, 2017.

[4] F. Rahimi, A. Ipakchi, and F. Fletcher, "The changing electrical landscape: end-to-end power system operation under the transactive energy paradigm," *IEEE Power and Energy Magazine*, vol. 14, no. 3, pp. 52–62, 2016.

[5] E. Münsing, J. Mather, and S. Moura, "Blockchains for decentralized optimization of energy resources in microgrid networks," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 2164–2171, 2017.

[6] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system." Available at https://bitcoin.org/bitcoin.pdf, last revised: 2008.

[7] C. Burger, A. Kuhlmann, P. Richard, and J. Weinmann, "Blockchain in the energy transition," *A survey among decision-makers in the German energy industry, Deutsche Energie-Agentur GmbH (dena), Germany*, 2016.

[8] E. Mengelkamp, J. Gärttner, K. Rock, S. Kessler, L. Orsini, and C. Weinhardt, "Designing microgrid energy markets: A case study: The brooklyn microgrid," *Applied Energy*, vol. 210, pp. 870–880, 2018.

[9] J. Luo, W. Su, and A. Q. Huang, "Bit-energy: An innovative bitcoin-style distributed transactional model for a competitive electricity market," in *IEEE Power & Energy Society General Meeting*, pp. 1–5, 2017.

[10] A. Hahn, R. Singh, C.-C. Liu, and S. Chen, "Smart contract-based campus demonstration of decentralized transactive energy auctions," in *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1–5, 2017.

[11] J. Wang, Q. Wang, N. Zhou, and Y. Chi, "A novel electricity transaction mode of microgrids based on blockchain and continuous double auction," *Energies*, vol. 10, no. 12, p. 1971, 2017.

[12] R. Gummadi, R. Johari, S. Schmit, and J. Y. Yu, "Mean field analysis of multi-armed bandit games." Available at SSRN: https://ssrn.com/abstract=2045842orhttp://dx.doi.org/10.2139/ssrn.2045842, last revised: August 11, 2016.

[13] D. Friedman, *The double auction market: institutions, theories, and evidence*. Routledge, 2018.

[14] P. Huang, A. Scheller-Wolf, and K. Sycara, "Design of a multi–unit double auction e–market," *Computational Intelligence*, vol. 18, no. 4, pp. 596–617, 2002.

[15] J. Niu, K. Cai, S. Parsons, P. McBurney, and E. H. Gerding, "What the 2007 tac market design game tells us about effective auction mechanisms," *Autonomous Agents and Multi-Agent Systems*, vol. 21, no. 2, pp. 172–203, 2010.

[16] J. Niu and S. Parsons, "Maximizing matching in double-sided auctions." Available at arXiv.org: https://arxiv.org/abs/1304.3135, last revised: Feb 11, 2013.

[17] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.

[18] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The nonstochastic multiarmed bandit problem," *SIAM journal on computing*, vol. 32, no. 1, pp. 48–77, 2002.

[19] Q. Peng and S. H. Low, "Optimal branch exchange for feeder reconfiguration in distribution networks," in *IEEE 52nd Annual Conference on Decision and Control (CDC)*, pp. 2960–2965, 2013.