An Online Learning based Optimal Bidding Approach for FTR Market Participants

Guibin Chen, *Student Member, IEEE*, Ye Guo, S*enior Member, IEEE,* Wenjun Tang, *Member, IEEE,* Qinglai Guo, *Senior Member, IEEE*, Hongbin Sun, *Fellow, IEEE,* and Wenqi Huang

***Abstract*—We consider the problem of optimal bidding and portfolio optimization for bidders in financial transmission rights (FTR) auction market. Based on the price-taker assumption, each FTR market participant aims to maximize his profit, which is the difference between the clearing price and FTR revenue. However, both the clearing price and the FTR revenue, are random and unknown. An online learning methodology is proposed to learn the optimal bidding through updating its policy with the newest observations of clearing results. With bidding prices derived by the online learning algorithm, a budget constrained portfolio optimization problem is solved to distribute the budget among profitable FTRs. Compared to other state-of-art online learning approaches, our proposed tree-based bid searching (TBS) algorithm converges faster to the optimal bidding price and has favorable linearithmic time complexity.**

***Index Terms*—Financial transmission rights, online learning, congestions, conditional value at risk.**

I. Introduction

A. Backgrounds

THE growing penetration of renewable resources and the uncertain demand make the transmission system more likely to operate closer to its capacity[1]. Hence, the congestion in power system becomes less predictable and locational marginal prices (LMPs) are more volatile. The uncertainty of LMPs may make market participants exposed to unknown congestion price risks, which dampens the long-term power trading and thus cause inefficiencies of the market. In addition, the Independent System Operator (ISO) also expects a mechanism to effectively and reasonably allocate the collected congestion charges. All these desires facilitate the emergence of FTR, which is a risk-hedging instrument for mitigating congestion price risks for market participants and revenue redistribution mechanism for the ISO [2].

The FTR entitles its holders to be paid or charged at the value of congestion rents of specified transmission path collected by the ISO in the day-ahead market [3]. An FTR is specified by the transmission capacity, a directional nodes pair and certain holding period. The FTR has been considered as a successful mechanism and implemented in PJM, New York, New England, and other electricity markets for decades [4-7]. In general, there exist two categories of FTR: the FTR obligations and FTR options. Holders with FTR obligations are charged whenever the differences between LMPs of the source node and the sink node is negative. By contrast, the FTR options would never be a liability for holders in such situation [8].

The purpose of FTR is to provide insurance for traders with physical transaction in power market. Nevertheless, to hamper market power and facilitate liquidity, the ISO permits speculators without power transaction to purchase or sell FTRs. The use for arbitrage has become a common industry practice [9]. However, the distribution of LMPs and clearing prices remains unknown and even non-stationary. How to formulate bidding strategy to acquire maximum payoff from the FTR auction market becomes rather difficult for price-takers. Although a higher bidding price enlarges the probability to acquire FTR, it causes inefficiency of budget allocation and will encounter more financial loss whenever clearing price exceeds FTR revenue. Thus, to continuously arbitrage from FTR auction markets, a bidder needs to learn the optimal bidding from the interaction processes with the market.

B. Related work

The concept of FTR was first proposed by Hogan in 1992 [10]. Further, the mathematical framework of PJM's clearing mechanism for FTR auction was investigated by [11-13]. The AC-OPF based FTR auction model considering contingencies and system loss allocation was developed in [14] and [15]. The efficiency of implementing FTR in realistic power markets was analyzed in [16-19]. Holding FTRs also affects market participants' bidding behavior in energy market. [20] analyzed the impacts of FTR on load serving entities' bidding strategy. While how virtual traders manipulate day ahead LMPs to profit from FTRs was discussed in [21]. A comprehensive survey on FTR's effects on power system and evolved process was conducted in [22].

Literatures studied the bidding strategy and portfolio construction are rare. [23] establishes a bi-level optimization framework to solve the optimal bidding problem in FTR market. The advantage of such methodology lies in its ability to model the self-interested bidding strategies of the bidder (modeled as the upper-level optimization problem - UL) and the clearing process of the ISO (modeled as the lower-level – LL). However, there exist inevitable limitations on such model: solving the LL problem requires the knowledge of sensitivity parameters of the network and bidding decisions of competitors. Since the FTR auction market acts as an opening market with hundreds of FTR

This work was supported in part by the National Key R&D Program of China (2020YFB090600, 2020YFB0906005).

G. Chen, Y. Guo (corresponding author), and W. Tang are with Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, Guangdong, China (email: guo-ye@sz.tsinghua.edu.cn).

Q. Guo and H. Sun are with the Department of Electrical Engineering, Tsinghua University, Beijing, China.

W. Huang is with the Digital Grid Research Institute, China Southern Power Grid, Guangzhou, China

products, a bidder has no access to system parameters and all potential competitors' bidding decisions. Another approach to solve the optimal bidding problem in FTR market falls into the category of reinforcement learning field [24][25]. However, previous work adopting reinforcement learning algorithm is based on the Q-learning method, which mainly relies on the Q-table to describe the action-value function. Due to discretization of continuous space, such method may suffer from the curse of dimensionality.

The effect of distribution factors in financial transmission right application was investigated by [26]. Based on the conclusion of [26], an approach for optimal FTR portfolio construction focuses on identifying congested transmission lines in the network is presented in [27]. With the information of *power transfer distribution factors* (PTDF), market players can select FTRs impacted significantly by the congested lines. Nevertheless, market players have no access to information of PTDF of the entire system. The approach proposed in [27] also assumes a fixed topology, which generally constitutes a limited assumption. Besides, the strategy to obtain FTRs in auction market is not considered in [28].

C. Scope and Contributions

The goal of this work is to develop a bidding strategy for FTR bidder based on online learning method. Through bidding in FTR auction, the bidder aims to determine bidding prices for available FTRs and allocate his budget among them to minimize risk related to the portfolio. By online learning we mean that the policy that determines optimal bidding price for each FTR is dynamically updated with newly observed information. In particular, we consider the objective of the online learning policy as maximizing expected cumulative unit profit of FTR.

More specially, the main contributions of this work are summarized as follows:

The optimal bidding problem in FTR auction is solved using an online learning scheme, which is based on proposed TBS algorithm. Unlike previous works solving optimal bidding problem in FTR auction with complete information, our work solves the examined problem with only instantaneous clearing outcomes as input and provides a practical approach for FTR bidders.

A budget constrained risk-minimization portfolio construction model is developed based on conditional value at risk (CVaR). Compared with previous online learning-based portfolio construction model [28], our work further includes purchase quantities as decision variables, which makes the proposed model more efficient in portfolio optimization.

The proposed TBS algorithm deals directly with continuous space and has a strong convergence property over finite auction rounds. In addition, TBS algorithm has a favorable computational complexity as $O(n \log n)$, making it an effective methodology for big data scenarios.

The reminder of this paper is organized as follows: Section II formulates the overall framework. Section III details the proposed online learning methodology. The performance of the proposed TBS algorithm and the portfolio optimization model are discussed in Section IV. Section V presents numerical results. Conclusions are discussed in Section VI.

II. Problem Modeling

A. Preliminary

FTRs are issued by the ISO through the auction market. Any bidder is required to submit his bidding price and purchasing *MW* amount for specified FTR in the auction market. Then ISO conducts a clearing process with collected bidding prices from all bidders under the constraints of transmission capacity and contingencies. Meanwhile, to ensure that collected congestion rents are higher or at least equal to the payments paid to FTR holders, ISO conducts simultaneous feasible test (SFT) in clearing process. ISO entitles bidders whose bidding prices are no less than the clearing prices with the specified FTRs for certain period and winning bidders pay the clearing prices as the cost.

In this paper, we take the PJM FTR auction market as example, which consists of four types of markets: long-term, annual, monthly, and secondary market. Both the long-term and the annual FTR auction are multi-round auction. While monthly FTR auction is a single-round auction. FTRs acquired in the monthly FTR auction have one-month term and the auction would be started approximately half a month ahead for FTRs awarded in next month. The PJM FTR secondary market is a bilateral trading system in which market participants could directly trade with each other. Without loss of generality, only the monthly FTR auction would be discussed.

The hourly economic value of FTR obligation is calculated as

$$P = (\lambda_{\text{sink}} - \lambda_{\text{source}}) * Q \tag{1}$$

where

λ_{sink} LMP of sink node in day ahead market

λ_{source} LMP of source node in day ahead market

Q *MW* amount of the FTR

As shown in equation (1), economic value of FTR obligation turns into a liability whenever λ_{source} is larger than λ_{sink}. In contrast, FTR option only become effective when economic value is beneficial to holders.

For a bidder, his profit for bidding an FTR obligation is

$$P_i = \left(\sum_{h=1}^{H} \Delta\lambda_{i,h} * Q_i - \pi_i * Q_i\right) \mathbb{1}\{b_i \geq \pi_i\} \tag{2}$$

where

H number of hours for holding an FTR obligation

P_i payoff of FTR i

$\Delta\lambda$ hourly LMP differences between pricing nodes of FTR *i* during the holding period

Q_i purchasing MW amount of FTR *i*

π_i market clearing price for FTR *i*

b_i bidding price for FTR *i*

For FTR option, the payoff for holding it is

$$P_i = \left(\sum_{h=1}^{H} \max(\Delta\lambda_{i,h}, 0) * Q_i - \pi_i * Q_i\right) \mathbb{1}\{b_i \geq \pi_i\} \tag{3}$$

different from FTR obligation, revenue of FTR option remains non-negative. Here the $\mathbb{1}\{.\}$ acts as an indicator function which equals to one if the argument is true otherwise zero. Without loss of generality, only the FTR obligations are discussed in this paper.

Next, we study the problem that an FTR bidder who intend to bid on M FTRs aims to identify the optimal b_i for each $i \in \{1,2, \ldots M\}$ and construct optimal portfolio under a certain budget constrained.

B. Bi-level optimization

In this sub-section, we develop a bi-level optimization model for repeated auction in FTR monthly market. At each round, an FTR bidder needs to decide optimal bidding price for each FTR and determine how to allocate his budget. Therefore, the optimal bidding problem and portfolio optimization problem need to be solved separately, which formulates a bi-level optimization problem.

C. Lower-Level Problem

The lower-level problem focuses on exploring each FTR's optimal bidding that maximizes the expected cumulative unit profit during N rounds of FTR monthly auction (8). Since the joint distribution of future FTR revenue and clearing prices remains unknown, it is impossible to solve the lower-level problem analytically:

$$\max_{b_i} \mathbb{E}\left(\sum_{n=1}^{N}\left(\sum_{h=1}^{H}\Delta\lambda_{i,h,n}-\pi_{i,n}\right)\mathbb{1}\left\{b_i \geq \pi_{i,n}\right\}\right) \tag{4}$$

An alternative way is to formulate policy μ to determine bidding price that maximizes expected cumulative unit profit based on newly observed clearing results $\{\sum_{h=1}^{H}\Delta\lambda_{i,h,n-1},\pi_{i,n-1}\}$and previous outcomes since the revenue and clearing price for FTR i of last round of auction can be observed at the beginning of each auction. Denote the policy μ as a sequence: $\mu_0, \mu_1, \ldots \mu_{N-1}$, which is updated with the increase of auctions. So, the bidding price b_i for FTR i is updated by the newly updated policy μ_n.

We develop an online learning algorithm to act as the policy μ. Based on the newly observed clearing results, the policy is updated sequentially and adaptively, which dynamically tracks the optimal bidding price.

D. Upper-Level Problem

The upper-level problem is a budget constrained portfolio optimization problem as shown in Eq. (5)-(8).

$$\min(CVaR_\alpha(\sum_{i=1}^{M} f\left(Q_{i,n},R_{i,n}\right)) \tag{5}$$

subject to

$$Q_{i,n} \geq 0 \tag{6}$$

$$Q_{i,n} * b_{i,n} \leq B \tag{7}$$

$$f\left(Q_{i,n},R_{i,n}\right) = -Q_{i,n} * R_{i,n} \tag{8}$$

where $i \in \{1,2, \ldots M\}$.

Based on the bidding prices $b_{i,n}$ derived from lower-level problem for each $i \in \{1,2, \ldots M\}$ at round n, the upper-level problem aims at minimizing risk associated with decision variable $Q_{i,n}$, which represents the purchased quantity for FTR i at round n. The objective function is defined as the conditional value at risk (CVaR) with confidence level α. Eq. (7) is the budget constraint, where the B represents a fixed budget for each round. The loss function $\sum_{i=1}^{M} f(Q_{i,n},R_{i,n})$ denotes the summation of the loss of individual FTR bidding (Eq. (8)), where $R_{i,n}$ represents the potential unit profit for holding FTR i at round n. The details about the CVaR instrument is discussed in Section IV-B.

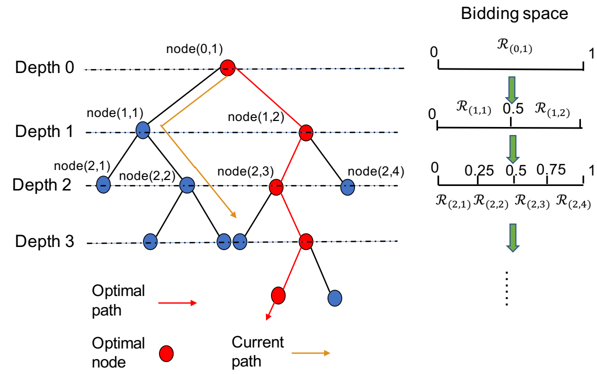

Fig. 1. Illustration of optimal bid search in the tree.

III. Online Learning Approach for FTR Bidding

In this section, we propose a TBS algorithm to solve the lower-level problem in Section II-C. The proposed methodology is inspired by the binary tree structure in [29], which deals with online optimization problem using a binary tree. Nodes of the tree are correlated to measurable regions of the action space. The TBS algorithm utilizes nodes of the binary tree to cover the corresponding size of bidding space and the tree is constructed incrementally.

The bidding space is explored under the guidance of upper bounds on the potential profit of bids covered within different nodes. One key insight is that the tree is expanded by refining optimistic nodes to increase the resolution of nodes as well as the associated bidding space to find more accurate bids, which enables to approximate the global optimal bid with increase of rounds.

In each round $n = 0,1,2, \ldots N$, a bidder repeatedly submits a bidding price $b_{i,n}$ to ISO from the bidding space, a continuous interval, for purchasing FTR i of the next month. For simplicity, here the payoff is normalized and bounded in continuous interval [-1,1] while the bidding price and clearing price is bounded in continuous interval [0,1].

Denote the entire bidding space by $\mathcal{B}$, which consists of bids from the continuous interval [0,1]. To describe the similarity between different bids, we assume there exists a dissimilarity function $\mathcal{M}$ which is a non-negative mapping $\mathcal{M}:\mathcal{B}^2 \to \mathbb{R}$. For example, $\mathcal{M}(b,b') > 0$ and $\mathcal{M}(b,b') = 0$ for all $(b,b') \in \mathcal{B}^2$. Noted that the dissimilarity $\mathcal{M}$ is only used for theoretical analysis and the algorithm has no requirement for an explicit input of $\mathcal{M}$. Generally, the dissimilarity could be some positive power of the Euclidean norm.

Given the concept of the dissimilarity function $\mathcal{M}$, the size of $\mathcal{B}$ and its subsets could be measured by $\mathcal{M}$ and expressly defined as

$$\text{Diam}(\mathcal{B}) = \max_{b,b' \in \mathcal{B}} \mathcal{M}\left(b,b'\right) \tag{9}$$

Fig. 1 presents the tree structure as well as partitioning process of bidding space. Each node (d,i) covers a particular bidding space $R_{d,i}$ (i.e., the root node which covers $\mathcal{B}$ is indexed by (0,1)), where d denotes the depth of the node and i denotes the rank of nodes with the same depth. For node (d,i),

$(d+1,2i-1)$ and $(d+1,2i)$ represent its two children nodes, which have no intersection with each other:

$$\mathcal{R}_{0,1} = \mathcal{B} \tag{10}$$

$$\mathcal{R}_{d,i} = \mathcal{R}_{d+1,2i-1} \cup \mathcal{R}_{d+1,2i} (\forall d \geq 0 \text{ and } 1 \leq i \leq 2^h) \tag{11}$$

$$\mathcal{R}_{d+1,2i-1} \cap \mathcal{R}_{d+1,2i} = \varnothing \tag{12}$$

As shown above, the covering bidding space monotonically decreases with the increase of depth, which implies narrowing the searching range to explore more accurate solution. Meanwhile, a reasonable assumption is made to efficiently denote the size of nodes in different depth.

Assumption 1. *For all nodes and their covering bidding space, given constants $L > 0$ and $m \in (0,1)$, there exists dissimilarity function $\mathcal{M}$ such that for any node (d,i):*

$$\text{Diam}\left(\mathcal{R}_{d,i}\right) \leq L(m)^d \tag{13}$$

Based on above assumption, the size of the bidding sub-space covered by node (d,i) is bounded by $L(m)^d$. As the number of nodes increase exponentially with the increase of depth d, it is reasonable to use the exponential decreasing term $L(m)^d$ to bound the size of covering regions. Noted that the $Diam$ is based on the dissimilarity, which can be adjusted by selecting appropriate mapping given constants L and m.

In addition, we assume that the minimum bidding space's size in each depth is bounded by $L_1(m)^d$. For notation simplicity, here we use $\mu(b) = \mathbb{E}(\frac{1}{n-1}\sum_{n'=1}^{n-1}(\sum_{h=1}^{H}\Delta\lambda_{n',h} - \pi_{n'})\mathbb{1}\{b \geq \pi_{n'}\})$ to denote the expected mean payoff that could be obtained by a fixed bidding price b up to round n. It is reasonable to believe that an individual has negligible influence on the market, therefore similar bidding prices from a price-taker could acquire similar mean payoff with randomly distributed clearing prices and FTR revenue. We assume that for bidding prices from the same node covering space, the differences of their mean payoff can be upper bounded by the dissimilarity $\mathcal{M}$.

Assumption 2. *For bids (b,b') from same node, the differences of their expected payoff could be upper bounded by:* $\mu(b) - \mu(b') \leq \mathcal{M}(b,b')$.

Noted that above assumption is appropriate for bids that are close. Due to the randomly distribution of payoff, two close bids may acquire widely varying instantaneous payoff in practical. Nevertheless, when it comes to the empirical mean payoff, two close bids would achieve similar expected payoff. For the convenience of mathematical analysis, the number for near-optimal nodes requires to be upper bounded. The number of β'-cover ball within a β-optimal node could be bounded by $C_h(\beta')^{-d_c}$, where d_c is the near-optimality dimension defined in *Definition 5* of [24], $\beta > \beta'$ and C_h is a constant.

Due to space limit, details of TBS algorithm's implementation procedure are discussed in Appendix A in [38].

IV. Performance Assessment

A. Performance of TBS algorithm

1) Regret Analysis

As defined by previous online machine learning literatures, the performance of such algorithm is described by the regret, which is the differences between cumulative reward acquired by the algorithm's strategy and the best fixed price in hindsight over total periods [28-36]. The objective of the lower-level problem is to find the global optimal bid that maximizes the total payoff over N rounds, which is defined as

$$b^* = \text{argmax}_{b\in\mathcal{B}}\, \mathbb{E}\left(\sum_{n=1}^{N}(\Delta\lambda_n - \pi_n)\,\mathbb{1}\{b^* \geq \pi_n\}\right) \tag{14}$$

Then the regret is defined as the differences between the expected cumulative payoff acquired by the TBS's bidding policy and that of the optimal bid b^*, *i.e.*,

$$R_N = \mathbb{E}\left(\sum_{n=1}^{N}\left(\sum_{h=1}^{H}\Delta\lambda_{n,h} - \pi_n\right)\mathbb{1}\{b^* \geq \pi_n\}\right) - \mathbb{E}\left(\sum_{n=1}^{N}\left(\sum_{h=1}^{H}\Delta\lambda_{n,h} - \pi_n\right)\mathbb{1}\{b_n \geq \pi_n\}\right) \tag{15}$$

The regret is monotonically nondecreasing and grows linearly with N for the worst case. Since we define the optimality is to maximize expected payoff for holding an FTR, observe that the algorithm converges to the optimal solution if the incremental regret of (14) goes to zero as $n \to \infty$.

Several important remarks are presented for demonstrating that R_N of TBS grows sub-linearly with N.

Remark 1: Based on the threshold value of selected times T_d^* (see Appendix A) for nodes in depth d, the depth $\mathcal{D}(n)$ of the tree $\mathcal{T}_n$ is bounded as

$$\mathcal{D}(n) \leq \frac{\ln\left(nL^2 / c^2 m^2\right)}{2(1-m)} \tag{16}$$

Proof. See Appendix B in [38].

Remark 1 illustrates that the exploration and exploitation policy of TBS bounds the increase rate of tree's depth, which only increases logarithmically with rounds n. Such $\mathcal{O}(\ln n)$ growth rate reduces the computation complexity of the algorithm since it bounds the growth rate of nodes amount of the tree.

Further, the regret is considered in situations depended on whether the differences between empirical mean payoff and expected mean payoff of a node are within a confidence interval. Remark 2 as below illustrates the probability of corresponding situation.

Remark 2: The situation that differences between empirical mean payoff and expected mean payoff of node (d,i) are within a confidence interval is described as

$$\mathcal{H}_n = \left\{ \bar{p}_{d,i}(n) - \mu\left(b_{d,i}\right) \leq c\sqrt{\frac{\ln\frac{\sqrt{4N-2n}}{\varepsilon}}{T_{d,i}(n)}} \right\} \tag{17}$$

with $c = 3\sqrt{1/(1-m)}$, $\varepsilon = \sqrt[18]{m/L}$, the probability for event $\mathcal{H}_n$ is $1 - \frac{1}{24}n^{-7}$.

Proof. See Appendix C in [38].

Denote the complementary event of $\mathcal{H}_n$ by $\mathcal{H}_n^c$. *Theorem 1* presents the upper bound of regret in situation that the differences between empirical mean payoff and expected mean payoff exceed the upper bound value in *Remark 2*.

Theorem 1: With the identified value of c and ε given in *Remark 2*, the upper bound for regret of TBS algorithm under event $\mathcal{H}_n^c$ is

$$R_N^{\mathcal{H}^c} \le \sqrt{N} \tag{18}$$

with probability $1 - 1/(144N^3)$.

Proof. See Appendix D in [38].

Theorem 1 demonstrates that the regret grows logarithmically with the increase of auction rounds N in situation $\mathcal{H}_n^c$. Such sublinear growth rate indicates that even under such minor probability situation, the algorithm enables to converge to the optimal. The main concertation should be attached on regret under the highest probability situation $\mathcal{H}_\mathrm{n}$. *Theorem 2* below shows the regret.

Theorem 2: Under the situation $\mathcal{H}_n$, the upper bound for regret of TBS algorithm on total N rounds is

$$R_N^{\mathcal{H}} \le 24\sqrt{2}\left(\frac{L^{d_c}c^2C_hL_1^{-d_c}m^{d_c}}{1-m}\right)^{\frac{1}{d_c+2}}\log\left(\frac{2\sqrt{N}}{\varepsilon}\right)^{\frac{1}{d_c+2}}N^{\frac{d_c+1}{d_c+2}} \tag{19}$$

Proof. See Appendix E in [38].

As shown in *Theorem 2*, the regret grows sub-linearly as the increase of auction rounds under situation $\mathcal{H}_n$, which implies the incremental regret would converge to zero as $N \to \infty$. Therefore, the TBS algorithm enables to converge to the optimal bid with the increase of auction rounds. By *Theorem 1* and *Theorem 2*, we demonstrate that the proposed online learning algorithm is robust to converge to optimal bid under all possible situations. Through finite rounds of interactions with market clearing process, the proposed algorithm shall converge to the optimal bid, thus provides solution for the lower-level problem in Section II.

2) Computational Complexity

In addition to this main conclusion, the design of our TBS algorithm has a favorable linearithmic computational complexity as $O(n\log n)$(see Appendix F in [38] for detailed analysis). In contrast, computational complexity of state of art methods for continuous space like DPDS [23] and EXP-Tree [22] is $O(n^2)$. Apparently, our TBS algorithm is more suitable for big data scenarios.

B. Budget constrained portfolio optimization

The budget constrained portfolio optimization problem (Eq. (5)-(8)) established in Section II is a convex optimization problem. Since the objective function is to minimize the CVaR of portfolio, which is a convex function with respect to the purchasing quantities in FTR market. In this sub-section we will convert the optimization problem into a linear programming problem.

For all Q_i where $i \in \{1,2,\dots M\}$, the portfolio loss $\sum_{i=1}^M f(Q_i, R_i)$ is a random variable having a distribution over $\mathbb{R}$ because of uncertain R_i. For convenience, we use vector $\boldsymbol{R}$ to denote R_i for all FTRs. The underlying probability distribution of $\boldsymbol{R}$ is represented by $p(\boldsymbol{R})$. Therefore, the probability that the portfolio loss $\sum_{i=1}^M f(Q_{i,n}, R_{i,n})$ does not exceed a value δ is defined as

$$\psi(Q_i,\delta) = \int_{\sum_{i=1}^M f(Q_i,R_i)\le\delta} p(\mathbf{R})d\mathbf{R} \tag{20}$$

Under a certain confidence level α, the value at risk (VaR) is defined as the minimum δ that is greater than the portfolio loss with probability α.

$$VaR_\alpha = \min\{\delta : \psi(Q_i,\delta) > \alpha\} \tag{21}$$

Then the CVaR of the portfolio under a certain confidence level α is defined as

$$\begin{aligned} &CVaR_\alpha\left(\sum_{i=1}^M f(Q_i,R_i)\right) \\ &=(1-\alpha)^{-1}\int_{\sum_{i=1}^M f(Q_i,R_i)\ge VaR_\alpha}\sum_{i=1}^M f(Q_i,R_i) * p(\mathbf{R})d\mathbf{R} \end{aligned} \tag{22}$$

The function $F_\alpha(Q_i, VaR_\alpha)$ has been defined as an upper bound of $CVaR_\alpha$, which is convex and continuously differentiable [25]

$$\begin{aligned} F_\alpha(Q_i,VaR_\alpha) = VaR_\alpha + (1-\alpha)^{-1}\int_{R_i}[\sum_{i=1}^M f(Q_{i,n},R_{i,n}) \\ -VaR_\alpha]^+ p(\mathbf{R})d\mathbf{R} \end{aligned} \tag{23}$$

An analytical expression of $p(\mathbf{R})$ is not required. By using Monte Carlo method to generate samples contain clearing results and FTR revenue, $F_\alpha(Q_i, VaR_\alpha)$ can be further simplified as

$$\begin{aligned} F_\alpha(Q_i,VaR_\alpha) = VaR_\alpha + \frac{1}{K(1-\alpha)} \\ \sum_{k=1}^K[\sum_{i=1}^M f(Q_{i,n},R_{i,n}) - VaR_\alpha]^+ \end{aligned} \tag{24}$$

where K denotes the number of selecting samples and $K = 2000$. Through introducing a substitution variable ζ_i^k for sample k of FTR i, $F_\alpha(Q_i, VaR_\alpha)$ is simplified as

$$F_\alpha(Q_i,VaR_\alpha) = VaR_\alpha + \frac{1}{K(1-\alpha)}\sum_{k=1}^K \zeta_i^k \tag{25}$$

subject to

$$\zeta_i^k \ge \sum_{i=1}^M f(Q_{i,n},R_{i,n}) - VaR_\alpha \tag{26}$$

$$\zeta_i^k \ge 0 \tag{27}$$

Therefore, the original budget constrained FTR portfolio optimization (6)-(9) for round n could be reformulated as

$$\min\left(F_\alpha(Q_i,VaR_\alpha)\right) \tag{28}$$

subject to

$$(6)-(8),(24),(25)$$

where $i \in \{1,2,\dots M\}$.

The above optimization problem could be efficiently solved by commercial solvers such as CPLEX and Gurobi.

V. Numerical Example

In this section, the proposed online learning methodology-based bidding strategy was applied to IEEE 14-bus system. To simulate practical electricity market condition, all loads are time-varying, which results in LMP fluctuations across the system. Based on the security constrained economic dispatch (SCED) simulation, we obtain 200 samples which contain FTR monthly revenues pair with corresponding market clearing prices, which fluctuates around the monthly revenue randomly. The proposed algorithm together with other online learning approaches is tested on offline dataset. Based on the IEEE 14-bus system, we formulate 14 FTRs constituted by nodes pairs across locations in the topology. For simplicity, the simultaneously feasibility test (SFT) won't be conducted to decide the issued FTR capacity. We assume that the bidder, as a price-taker whose demand merely occupies a small proportion of total available amount. Thus, the bidder can obtain the purchasing amount whenever his bid is determined as winning price. The first part presents that the expected unit profit of proposed TBS algorithm converges faster to that of optimal bid while compares to algorithms from relevant online learning literatures. Subsequently, the online learning methodology-based portfolio optimization model is compared with other portfolio construction methods, which are also based on online learning approaches.

A. Performance of TBS algorithm

1) Benchmarks

To demonstrate the performance of proposed TBS algorithm in converging to the global optimal bid, we compare it with two algorithms from recent advanced online learning literatures.

The first one is the EXP-Tree algorithm by [30], which was applied in repeated sequential auctions. In the sequential auction, in which auction item's value and bidding price are bounded in interval [0,1], EXP-Tree randomly submits a bidding price from [0,1] and observes the winning bid. Every round the winning bid would be embedded in the interval [0,1] recursively to derive sub-intervals, which generate submitted bidding prices for the auction. Then, the profit obtained by the bidding price would be used to update the weight of corresponding sub-interval. A sub-interval with higher weight has more opportunities to be selected.

The second one is the WIN-EXP algorithm proposed by [33], which aims to find the best fixed bidding price in hindsight for repeated sponsored search auctions. With a delicately formulated utility function, this algorithm updates weights distribution among finite bids set B. For example, the initialized weights distribution for each bid b in B is $w(b) = \frac{1}{|B|}$. With the formulation of utility function as $u_n(b) = \frac{(\Delta\lambda_n - \pi_n - 1)\mathbb{1}\{b \geq \pi_n\}}{\sum_{b' \in B, b' \geq \pi_n} w_n(b')}$, WIN-EXP updates corresponding bid option's weight as

$$w_{n+1}(b) = w_n(b) * \exp\left(\eta * u_n(b)\right) \tag{29}$$

Without loss of generality, the bidding interval [0,1] is uniformly discretized into 11 bidding options here. Since the number of simulated samples is only 200, a more intensive bid options partition would result in insufficient evaluation on bids set.

2) Results Analysis

We randomly select FTR pricing nodes pairs across locations in IEEE 14-bus system. Based on it, the monthly revenue and clearing price for each FTR has been simulated with 200 samples. For comparison among TBS, EXP-Tree and WIN-EXP algorithms, we arbitrarily select an FTR as the target in the auction market. Noted that the magnitude of FTR's unit revenue is normalized in [-1,1] and the bidding space is continuum interval [0,1] for EXP-Tree and TBS. For WIN-EXP, bids set is adaptively formulated as $[0, 0.1, 0.2, \ldots\ldots 0.9, 1]$. All these algorithms are applied in the repeated FTR monthly auction. To measure the performances of all these algorithms in converging to optimal solution, criterions like regret and average regret are used. As shown in Fig. 2, compared with EXP-Tree and WIN-EXP algorithms, TBS achieves a significant lower regret over total auction rounds.

Meanwhile, the WIN-EXP presents the highest regret magnitude in all periods. Such outcome indicates that a finite discretization bids set is inefficient to converge to global optimal. Fig. 3 shows the average regret of all these algorithms.

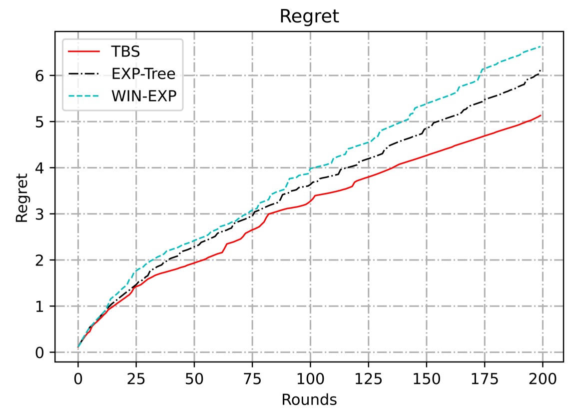

Fig. 2. Comparison of regret over 200 FTR monthly auction rounds.

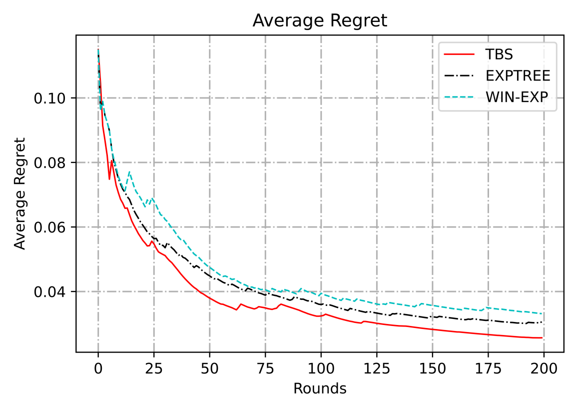

Fig. 3. Comparison of average regret.

Observed that the average regret curve of TBS is not as smoother as others, which is caused by the path re-selecting. With the accumulative of samples, all algorithms' average regret goes to a smooth descent stage. TBS outperforms WIN-EXP and EXP-Tree on average regret with the lowest average regret throughout the entire auction rounds.

B. Budget Constrained Portfolio Optimization

As illustrated above, compared with other advanced online learning methods, the expected profit of TBS converges faster to expected payoff of optimal bid. With the derived bidding prices by the online learning algorithm, the optimized FTR portfolio construction model in UL problem (5)-(8) further decides purchasing *MW* amount for each FTR. In this subsection, performance of such portfolio construction model would be presented.

1) Benchmark Methods

We compare the proposed portfolio construction model with other online learning method-based portfolio optimization models for evaluating its performance in practical application. The first one is the DPDS model from [28]. Similar to our work, [28] develops an online learning algorithm to find global optimal bid in sequential virtual trading. Based on bidding prices derived by the online learning algorithm for all options, DPDS adopts dynamic programming to accomplish the optimal portfolio construction. In accordance with procedure of DPDS, the bids set is generated by recursively discretizing the continuum interval $[0,\mathcal{B}]$, where $\mathcal{B}$ is the total budget for each round. For example, DPDS's bids set is $\mathcal{F}_n = [0, \frac{\mathcal{B}}{\alpha_n}, \frac{2\mathcal{B}}{\alpha_n}, \dots\dots \frac{(\alpha_n-1)\mathcal{B}}{\alpha_n}, \mathcal{B}]$, where $\alpha_n = n+1$. Then, the empirical average payoff of bid in bids set for FTR k would be calculated as

$$\bar{p}_{k,n}(b) = \frac{1}{n-1}\sum_{s=1}^{n-1}(\Delta\lambda_s - \pi_s)\,\mathbb{1}\{b \geq \pi_s\} \tag{30}$$

where $b \in \mathcal{F}_n$.

For each FTR, the bidding price that earns the highest empirical average payoff will be submitted for this round. Then with dynamic programming, DPDS will allocate budget for purpose of maximizing total expected payoff.

The other portfolio construction model based on online learning method is the UCBID-GR developed by [28] and derived from UCBID [30]. For each round of auction, the UCBID-GR observes revenue of each FTR and submits bid as the average FTR revenue up to current round. Then, according to the observed FTR revenues and corresponding clearing prices distributions, UCBID-GR would allocate budget to purchase the most profitable FTR until there is no sufficient residual budget.

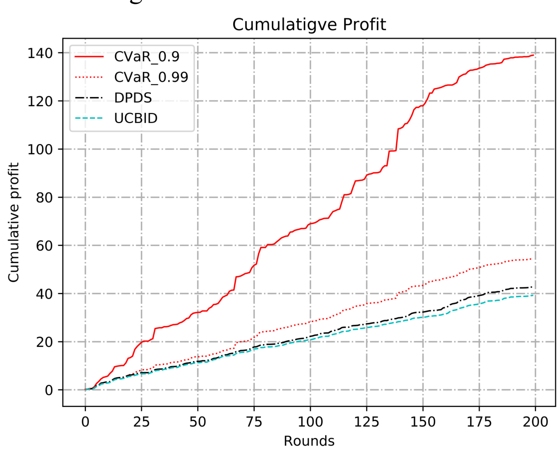

Fig. 4. Cumulative profit trajectory over 200 FTR auction rounds with $\mathcal{B} = 6$.

2) *Results Analysis*

Based on the IEEE 14-bus system, we obtain 200 samples for 14 FTR candidates' monthly revenue and clearing prices to simulate for 200 auction rounds' clearing outcomes. Similarly, each portfolio construction model is utilized to allocate same budget among all these FTR candidates based on bidding prices generated by corresponding learning algorithms. Noted that for simplicity, the magnitude range of FTR's unit monthly revenue is normalized in [-1,1] and the clearing price is normalized in [0,1].

The trajectory of cumulative payoff obtained by each model over total 200 months' auction with budget level $\mathcal{B} = 6$ is presented in Fig. 4. The inherent risk averse of CVaR model enables to make trade-off on profitability and risk on FTR auction. Observed from Fig. 4 that the CVaR model with confidence level at 0.9 achieves to obtain a higher cumulative profit over entire auction rounds. While the other CVaR model at 0.99 confidence level appears to be more smoother compare to CVaR(0.9). Since higher confidence level leads to a more risk averse behavior, which restrains to purchase risky FTR products and therefore reduce opportunities of acquiring more profit.

Unlike DPDS and UCBID-GR, which purchase profitable options for per unit amount at most. Our model enables to decide flexible purchasing quantities in the auction market. Observed that these two benchmark models, DPDS and UCBID-GR, could not collect more profit with higher budget. Similar to the numerical results shown in [28], the DPDS model outperforms the UCBID-GR by its ERM (empirical risk minimization) approach and dynamic programming (DP) method in portfolio construction. However, the DPDS considers the portfolio optimization as 0-1 Knapsack problem, which restricts to include quantities as decision variables. Therefore, even with higher budget level, these two models could not achieve significant profit increase. Obviously that the CVaR-based model is more compatible with practical situation.

The Sharpe ratio [1] is commonly used to gauge the performance of an investment [37]. The higher the ratio, the greater the investment return relative to the amount of risk taken. As presented in Table I, the DPDS and UCBID-GR have higher Sharpe ratio compared to CVaR(0.9) and CVaR(0.99), although these two models achieve higher cumulative profit in FTR monthly auction, their Sharpe ratio is less than DPDS and UCBID-GR under less budget level.

TABLE I
SHARPE RATIO OF PORTFOLIO CONSTRUCTION MODEL

Model	$\mathcal{B} = 1$	$\mathcal{B} = 6$	$\mathcal{B} = 14$
CVaR(0.9)	7.424	9.081	10.239
CVaR(0.99)	12.263	12.143	15.071
DPDS	17.129	16.358	16.362
UCBID-GR	17.474	17.474	17.474

As mentioned previously, DPDS and UCBID-GR just purchase per unit quantity for each profitable FTR product, which enable them to achieve higher Sharpe ratio as they minimize deviation of the return. In addition, the CVaR-based model includes quantity as decision variable and therefore lead to significant deviation of return. Especially for CVaR(0.9), which intends to allocate more budget for products with high profitability while focuses less on the risk of payoff fluctuation. Besides, as shown in Table I, the Sharpe ratio of DPDS and UCBID-GR remains the same when the budget level is sufficient for them to purchase profitable products. Even though there is still surplus in budget, no more quanties would be bought.

In contrast, Both CVaR(0.9) and CVaR(0.99) models achieve higher Sharpe ratio with the increase budget levels. As shown in Table I, with the increase budget level, the Sharpe ratio of CVaR(0.99) almost approximates the one of DPDS. The phenomenon indicates our CVaR-based model could also presents satisfied Sharpe ratio with sufficient budget. Such phenomenon illustrates that the proposed model is more compatible with realistic situation compared to the similar DPDS model. Since any speculators in practical FTR auction market may with budget level that is sufficient to purchase a

[1]In this paper, Sharpe ratio is calculated as $\sqrt{N}\frac{\overline{p_N}}{\sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(p_n-\overline{p_N})^2}}$, where $\overline{p_N} = \frac{1}{N}\sum_{n=1}^{N} p_n$. N is the total auction rounds under consideration and p_n is the profit of round n.

certain quantity of products instead of per unit amount.

VI. Conclusion

We address the optimal bidding problem and portfolio optimization problem in repeated FTR auctions, where bidders do not know the underlying profit of FTRs. We develop an online learning approach with only observed clearing outcomes as input to search optimal bidding price. The proposed TBS algorithm sets up the bidding problem in continuous space and can converge to optimal solution as auction rounds increase. Meanwhile, the well-designed binary tree expansion scheme realizes a favorable computational complexity as $\mathcal{O}(n\log n)$, making the algorithm an effective method for big data scenarios. Instead of simply considering bidding on a specific FTR, we combine the online learning algorithm with conditional value at risk (CVaR) tool, which can assess many FTRs simultaneously and derive optimal purchasing quantities among them. To summarize, we provide an effective methodology for an FTR bidder to search optimal bidding price and to analytically optimize his FTR portfolio with adjustable risk preference.

References

[1] T. Y. Hu, W. C. Wu, Q. L. Guo, H. B. Sun, L. B. Shi, X. W. Shen, "Very Short-term Spatial and Temporal Wind Power Forecasting: A Deep Learning Approach," *CSEE Journal of Power and Energy System*, vol. 6, no. 2, pp.434-443, 2020.

[2] O. Alsac, J. Bright, S. Brignone, M. Prais, C. Silva, B. Stott, and N. Vempati, "The rights to fight price volatility," *IEEE Power and Energy Mag*, vol. 3, no. 6, pp. 47–57, Jul–Aug. 2004.

[3] PJM, "Financial transmission rights," Audubon, PA, USA, 2020. [Online]. Available: https://www.pjm.com/-/media/documents/manuals/m06.ashx

[4] X. Ma, D. I. Sun, G. W. Rosenwald, and A. L. Ott, "Advanced financial transmission rights in the PJM market," in *Proc. IEEE/Power Eng. Soc. General Meeting*, Jul. 13–17, 2003, vol. 2, pp. 1031–1038.

[5] A. Bykhovsky, D. A. James, and C. A. Hanson, "The introduction of option financial transmission rights into the New England market," in *Proc. IEEE/Power Eng. Soc. General Meeting*, Jun. 12–16, 2005, pp. 284–289.

[6] K. W. Cheung, "Standard market design for ISO New England wholesale electricity market: An overview," in *2004 Proc. IEEE Int. Conf. Electric Utility Deregulation, Restructuring and Power Technologies (DRPT 2004)*, Apr. 2004, vol. 1, pp. 38–43.

[7] Midwest ISO Market Concept Study Guide. [Online]. Available: http://www.midwestmarket.org/Documents/EducationsMaterial/General Guides.

[8] S. Harvey, W. Hogan, and S. Pope, "Transmission capacity reservations and transmission congestion contracts," *Tech. Rep.*, [Online] Available: http://www.ksg.harvard.edu/whogan, 2005.

[9] O. Alsac, J. Bright, S. Brignone, M. Prais, C. Silva, B. Stott, and N. Vempati, "The rights to fight price volatility," *IEEE Power and Energy Mag*, vol. 3, no. 6, pp. 47–57, Jul.–Aug. 2004.

[10] W. W. Hogan, "Contract networks for electric power transmission," *J. Regulat. Econ,* vol. 4, no. 3, pp. 211–242, Sep. 1992, [Online] Available: https://link.springer.com/content/pdf/10.1007/BF00133621.pdf

[11] X. Ma, D. Sun, and A. Ott, "Implementation of the PJM financial transmission rights auction market system," in *Proc. IEEE Power Engineering Society Summer Meeting*, Jul. 2002, vol. 3, pp. 1360–1365.

[12] V. Sarkar and S. A. Khaparde, "A robust mathematical framework for managing simultaneous feasibility condition in financial transmission rights auction," in *2006 IEEE Power Engineering Society General Meeting*, Montreal, Que., 2006, pp. 6

[13] P. Ermida, J. Ferreira, Z. Vale and T. Sousa, "Auction of Financial Transmission Rights in electricity market environment," in *2010 7th International Conference on the European Energy Market*, Madrid, 2010, pp. 1-6.

[14] H. Xu, Y. Tang and Q. Wan, "Implementation of AC Optimal Power Flow Based Financial Transmission Right Auction under Static Security Constraints," in *2009 Asia-Pacific Power and Energy Engineering Conference*, Wuhan, 2009, pp. 1-4.

[15] A. Alburidy and L. Fan, "Loss allocation in AC OPF-based financial transmission rights auction models," in *2017 North American Power Symposium (NAPS)*, Morgantown, WV, 2017, pp. 1-6.

[16] S. Adamson, T. Noe, and G. Parker, "Efficiency of financial transmission rights markets in centrally coordinated periodic auctions," *Energy Economics*, 2010(32): 771-778.

[17] S. J Deng, S. Oren, and A.P. Meliopoulos, "The inherent inefficiency of simultaneously feasible financial transmission rights auctions," *Energy Economics,* 2010(32): 779-785.

[18] F. Kunz, K. Neuhoff, J. Rosellón, "FTR allocations to ease transition to nodal pricing: An application to the German power system," *Energy Economics,* 2016(60): 176-185.

[19] D. Molzahn and C. Singletary, "An empirical investigation of speculation in the MISO financial transmission rights auction market," *The Electrical Journal*, 2011(24): 57-68.

[20] X. Fang, F. Li, Q. Hu, Y. Wei, and N. Gao. "The impact of FTR on LSE's strategic bidding considering coupon-based demand response," in *2015 IEEE Power and Energy Society General Meeting,* Denver, CO, 2015, pp. 26-30.

[21] C. L. Prete, N. Guo, and U. V. Shanbhag, "Virtual bidding and financial transmission rights: An equilibrium model for cross-product manipulation in electricity markets," *IEEE Trans. Power Syst*, vol. 34, no. 2, pp. 953-967, 2019.

[22] V. Sarkar and S. A. Khaparde, "A Comprehensive Assessment of the Evolution of Financial Transmission Rights," *IEEE Trans. Power Syst*, vol. 23, no. 4, pp. 1783-1795, 2008.

[23] Tao Li and M. Shahidehpour, "Risk-constrained FTR bidding strategy in transmission markets," *IEEE Trans. Power Syst*., vol. 20, no. 2, pp. 1014-1021, 2005.

[24] N. P. Ziogos, A. C. Tellidou, V. P. Gountis and A. G. Bakirtzis, "A Reinforcement Learning Algorithm for Market Participants in FTR Auctions," in *2007 IEEE Lausanne Power Tech*, Lausanne, 2007, pp. 943-948.

[25] N.P. Ziogos, and A.C. Tellidou, "An agent-based FTR auction simulator," *Electric Power Systems Research*, 2011(81), pp.1239-1246.

[26] Minghai Liu and G. Gross, "Role of distribution factors in congestion revenue rights applications," *IEEE Trans. Power Syst.,* vol. 19, no. 2, pp. 802-810, 2004.

[27] D. Apostolopoulou, G. Gross and T. Güler, "Optimized FTR Portfolio Construction Based on the Identification of Congested Network Elements," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4968-4978, 2013.

[28] S. Baltaoglu, L. Tong and Q. Zhao, "Algorithmic Bidding for Virtual Trading in Electricity Markets," *IEEE Trans. Power Syst*., vol. 34, no. 1, pp. 535-543, 2019.

[29] S. Bubeck, R. Munos, G. Stoltz, and C. Szepesva ́ri, "X-armed bandits," *J. Mach. Learn. Res*., vol. 12, pp. 1655–1695, 2011.

[30] J. Weed, V. Perchet, and P. Rigollet, "Online learning in repeated auctions," in *29th Annu. Conf. Learning Theory,* 2016, pp. 1562–1583.

[31] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk, " *J. Risk*, vol. 2, pp. 21–42, 2000.

[32] M. G. Azar, A. Lazaric, and E. Brunskill, "Online Stochastic Optimization under Correlated Bandit Feedback," *Proc. Int. Conf. on Machine Learning (ICML)*, Beijing, pp. 1557-1565, 2014.

[33] Feng, Z., Podimata, C., and Syrgkanis, V, "Learning to Bid Without Knowing your Value," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018, pp.505-522.

[34] S. Bubeck, N. R. Devanur, Z. Huang and R. Niazadeh, "Online auctions and multi-scale online learning," in *Proceedings of the 2017 ACM Conference on Economics and Computation,* 2017, pp.497-514

[35] K. Amin, A. Rostamizadeh and U. Syed, "Learning prices for repeated auctions with strategic buyers," in *Proceedings of the 26th International Conference on Neural Information Processing Systems*, 2013, pp.1169-1177.

[36] M. Dudik, N. Haghtalab, H. Luo, R. E. Shapire, V. Syrgkanis and J. W. Vaughan. "Oracle-efficient online learning and auction design," in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, 2017, pp.528–539.

[37] W. F. Sharpe, "The Sharpe ratio," *J. Portfolio Manage*., vol. 21, no. 1, pp. 49–58, 1994.

[38] G.Chen, Y.Guo, W.Tang, Q.Guo, H.Sun, and W.Huang "Supplementary material for 'An online learning approach based trading strategy for FTR auction market'," [Online]. Available: https://github.com/GuibinChen92057/Supplementary-material