**Preprints of the 21st IFAC World Congress (Virtual)**
**Berlin, Germany, July 12-17, 2020**

# An Optimal Day-ahead Bidding Strategy and Operation for Battery Energy Storage System by Reinforcement Learning

**Yi Dong** * **Tianqiao Zhao** ** **Zhengtao Ding** *

* *Department of Electrical and Electronic Engineering, the University of Manchester, M13 9PL, Manchester, UK (e-mail: yi.dong@manchester.ac.uk, zhengtao.ding@manchester.ac.uk).*
** *Department of Electrical and Computer engineering, Southern Methodist University, PO Box 750100, Dallas, TX 75275, USA (e-mail: tianqiaoz@smu.edu)*

**Abstract:** The Battery Energy Storage System (BESS) plays an important role in the smart grid and the ancillary market offers high revenues. It is reasonable for the owner of the BESS to maximise their profits by deciding how to bid with their rivals and balance between the different market offers. Therefore, this paper proposes an optimal bidding model of the BESS to maximise the total profit from the Automation Generation Control (AGC) market and the energy market, while taking the charging/discharging losses and the life of the BESS into consideration. Taking advantages of function approximation approaches, a reinforcement learning algorithm is introduced to the designed model, which can cope with the continuous and massive states of the proposed model and avoid the dimension curse. The resultant novel bidding model would help the BESS owners to decide their biddings and operational schedules profitably. Several case studies illustrate the effectiveness and validity of the proposed model.

*Keywords:* Battery Energy Storage System (BESS), optimal bidding, reinforcement learning.

## 1. INTRODUCTION

The Battery Energy Storage System (BESS) will play an important role in the future smart grid. With the rapid development of battery technology, the BESS can bring more benefits for the owners, while its construction cost is gradually reduced (NEE (2018)). There will be more companies focusing on the development and construction of the BESS. As its capacity increases, the BESS will participate in different markets and benefit from multiple services (Michael et al. (2018)). Additionally, the frequency regulation market demands rapid response and offers high rewards (PJM (2011)), so that the BESS owners will put more attention on the regulation market with their BESS, which will lead to further competitions in the future smart grid. Therefore, how to allocate the capacity of the BESS and make bidding decisions have become important issues.

The technology of the BESS has been developed rapidly in the last decades. Some researchers built batteries with higher energy density, which could store more energy with the same volume (Albertus et al. (2018); Monti et al. (2019)). In the industry, some nickel, cobalt and manganese (NCM) 811 batteries were produced and applied to a part of the smart grid (Strehle et al. (2019); Coffin and Horowitz (2018)). Some researchers tried to extend the life of the battery, which can significantly reduce the batteries' costs. Liu et al. (2018b) studied the life of the widely used Li-ion battery. They used the lithium phosphorous oxynitride (LiPON) layer to extend at least 10 years of storage life and improve the capacity retention during storage ageing at elevated temperature. Geaney and O'Dwyer (2017) studied the fundamental operation and the stability of the Li-$O_2$ battery, which emphasised the importance of selecting proper discharge/charge rate and the discharge depth for other cathode/electrolyte combinations to improve the cycle life performance of Li-O2 batteries. More studies are related to the modelling and optimisation of the batteries. For example, Rashid and Gupta (2017) did some physical experiments on lithium batteries and put forward relevant numerical simulation models. Dong et al. (2019) studied the distributed battery optimisation strategies and applied in power system demand side management.

One major application for the BESS is frequency regulation services in the Automation Generation Control (AGC) market. Compared with other storage systems, the BESS has the advantages of easy storage, high reliability and fast response, which are more suitable for the frequency regulation market. Moreover, the AGC market offers 3 times mileages for RegD (Dynamic Regulation) service, which will bring high revenue for the BESS owners. As a result, more BESS owners were expected to compete in the AGC market and some researchers had been paying more attention to the AGC market (Xu et al. (2014a); Tan and Zhang (2017)). In Xu et al. (2014a), a control strategy for the BESS in frequency regulation was provided, considering the ageing cost while keeping the state of charge (SoC) of the BESS. In Tan and Zhang (2017), a coordinated control strategy of the BESS was proposed to ensure the wind power plants' commitment to frequency ancillary services, focusing on reducing the BESS's size and extending the lifetime of the BESS. However, these mentioned literature only considered the application of the BESS in one market. With the emergence of large-capacity BESS, some articles studied the operation strategies of the BESS in multiple markets, so as to maximise the overall profit of the BESS by controlling the placement proportion of the BESS in different markets. For example, He et al. (2016) integrated the energy storage system and solar power plant. They considered the energy, reserve and regulation market, and proposed an optimal operation strategy for the solar power plants.

Another problem missed by these literature is that the bidding strategies only solve the allocation problem of a single BESS, in which their bidding rivals are neglected. With the entry of the rivals, the bidding market of the BESS presents some challenges. During the process of bidding, the bidder does not know the rivals' bidding price and bidding quantity, which is hard to solve by traditional optimisation algorithms. Furthermore, since bidding is a highly random and uncertain process, the bidders cannot know the specific revenue model during bidding. They only know the market clearing results from the system operator (SO) in the smart grid. Fortunately, there were many types of bidding researches in other applications (Li et al. (2018); Nanduri and Das (2007)) and the effectiveness of the reinforcement learning algorithm in bidding problems were proved. Li et al. (2018) applied the model-free reinforcement learning algorithm to solve the optimal carbon capture problem in the wholesale market bidding case. Nanduri and Das (2007) formulated a stochastic game model for the energy market and proposed a reinforcement learning based solution methodology.

Therefore, this paper proposes a novel model that determines the optimal bidding strategy of a BESS in day-ahead energy and regulation markets, considering the charging/discharging losses and the ageing cost of the BESS. Additionally, the reinforcement learning algorithm is applied to the proposed model in order to solve the multiple rival bidding problem. At the same time, the function approximation approach is introduced in this paper to address the redundancy caused by massive data and therefore prevent the dimension curse. Based on the proposed model, the BESS could obtain a more accurate and profitable bidding strategy.

The major contributions of this paper are summarised as follows: First, the bidding model considered in this paper is more accurate and more applicable. Then, the proposed bidding strategy can help the BESS owner to win regulation offers without rivals' information. Moreover, this paper applied the function approximation tool to analyse the discrete-time bids states, which significantly reduced the training time.

## 2. MARKET DESIGN

This section studies the bidding mechanism of battery energy storage system in different power markets.

With the development of battery technology, the capacity of the BESS is increasing rapidly. According to the importance of batteries in AGC market service, we assume that the BESSs have the market power to influence AGC market (NEE (2018)). Since the main services and revenues

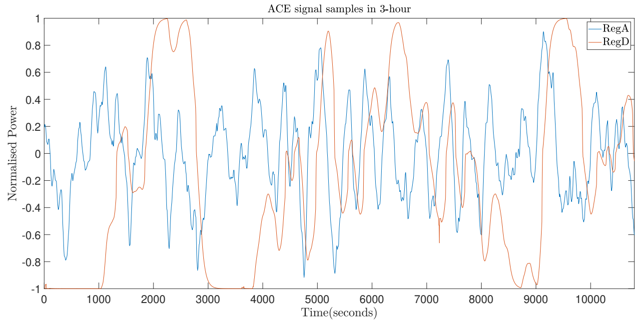


Fig. 1. Real-time RegD amd RegA data.

of the BESS come from the AGC market, according to Michael et al. (2018), supplying sufficient power and energy capacity for the AGC market has the highest priority among all the services from the perspective of the system operator. In this paper, based on the prediction of energy market and AGC market, the winning bids of the BESS are determined considering the AGC market conditions.

### *2.1 Automatic Generation Control (AGC) Market*

In the AGC market, the operation of smart grid must be subjected to keep the supply and demand balance. During the frequency control, the supply-demand balance of the whole network is met by adjusting the output of frequency modulation unit, the BESS.

The AGC market obtain the mismatched power by

$$\Delta P_{td} = P_{load} - P_{energy} - P_{plan} \tag{1}$$

where $\Delta P_{td}$ is the power mismatch; $P_{load}, P_{energy}$ and $P_{plan}$ are the load demand, renewable power output and the planned power output, respectively.

In power grid dispatching, area control error (ACE) are usually sent to AGC with a period of 2-4 seconds.

$$P_{ACE} = \Delta P_{td} + \beta_f \cdot \Delta f \tag{2}$$

where $\beta_f$ is the coefficient of frequency deviation, $\Delta f$ is the frequency deviation and $P_{ACE}$ is the ACE signal.

In the smart grid, ACE signal can be divided into a dynamic regulation signal (RegD) and a traditional regulation signal (RegA) by a low-pass filter, as shown in Fig. 1. According to He et al. (2015), the mileage ratio of RegD is approximately three times larger than RegA's mileage ratio. Thus the service of RegD can obtain higher revenue and compensation than RegA. Moreover, RegD has the characteristic of zero mean, which helps to reduce the requirement of the BESS capacity. Although the mean value of RegD is zero, due to the loss of charging and discharging, the SoC of the BESS will continue to decline unless additional power is injected to compensate the energy loss for a longer time running. To this end, this paper includes the charging/discharging loss model of the BESS, and the details will be introduced in the next section.

### *2.2 Energy Market*

The revenue of energy market is mainly from the planned output power. Compared with traditional generating units, a BESS only supplies or consumes small portion of electricity, the BESSs are supposed to be the price-takers, who will not affect the electricity price in the energy market. The BESS will submit the day-ahead bids to the energy market system operator, and then the system operator will allocate the electric energy according to different requirements. Since the BESS has the characteristics of low cost, good power quality and fast response, we assume that the battery will win the bids in the energy market. Therefore, the revenue in the energy market can be described as:

$$R_{e,t} = p_t \cdot b_{e,t} \tag{3}$$

where $p_t$ is the clearing price, $b_{e,t}$ is the energy bidding quantity of the BESS and $R_{e,t}$ is the revenue of the BESS in energy market at time slot $t$.

### *2.3 Model of BESS*

The BESS unit should provide AGC services frequently in long term running. Therefore, two types of the BESS costs are considered in this paper, i.e., charging/discharging loss cost and the BESS ageing cost.

*Loss Cost of BESS* According to Xu et al. (2014b), charging efficiency and discharge efficiency are different, and the efficiency of charging/discharging has an approximate linear relationship with charging/discharging power, which can be formulated as follows:

$$\eta_{charge} = \alpha_{charge} + \beta_{charge} \cdot P_{charge} \tag{4}$$

$$\eta_{discha} = \alpha_{discha} + \beta_{discha} \cdot P_{discha} \tag{5}$$

where $(\alpha_{charge}, \beta_{charge})$ and $(\alpha_{discha}, \beta_{discha})$ are the coefficients of the charging and discharging process. $\eta_{charge}$ and $\eta_{discha}$ are the charging and discharging efficiency, respectively. We assume that the electricity price is $p_t$. The charging/discharging losses then represented as

$$C_{chaloss} = p_t \cdot P_{charge}(1 - \eta_{charge}) \cdot \Delta T \tag{6}$$

$$C_{disloss} = p_t \cdot P_{discha}\left(\frac{1}{\eta_{discha}} - 1\right) \cdot \Delta T \tag{7}$$

where $\Delta T$ is the control period of regulation service and it is set as 4 seconds.

*Ageing Cost of BESS* Ageing cost is an important expenditure when the BESSs provide the power system service, and the BESS may not meet the requirements of the system after excessive ageing. Therefore, the ageing cost model needs to be considered when calculating the revenue of the BESS. Based on Zhou et al. (2017), the maximum energy capacity of the BESS will be reduced by the increase of charging/discharging cycles.

For different type of battery, $N_d^{fail}$ is a function of DOD(%), which can be calculated as

$$N_d^{fail} = N_{100}^{fail} \cdot d^{(-k_P)} \tag{8}$$

where $N_d^{fail}$ is the maximum number of charging/discharging cycles at $d$ DOD, $d$ is the depth of discharge (DoD), and $k_P$ is a constant parameter for different type of batteries ranging from 1.1 to 2.2 (Ying et al. (2016)). To calculate precisely, the equivalent 100%-DOD cycle number $n_{100}^{eq}$ of $n_d$ cycles at $d$ DoD is described as

$$n_{100}^{eq} = n_d \cdot d^{k_P} \tag{9}$$

In practice, the battery cannot predict the positive and negative power command signals to be given by the system operator, so that the BESS will provide one bid for charging and one bid for discharging at each time slot. Thus, the ageing cost of the BESS can be formulated as

$$C_{ag,t} = \frac{n_{100,t+1}^{eq} - n_{100,t}^{eq}}{2 \cdot N_{100}^{fail}} \cdot C_{inv} \tag{10}$$

Combining (9) and (10), we can obtain the equivalent ageing cost as

$$C_{ag} = \sum_{j=0}^{24*3600/\Delta T} n_d \frac{(d_j + \frac{P \cdot \Delta T}{E})^{k_P} - (d_j)^{k_P}}{2 \cdot N_{100}^{fail}} \cdot C_{inv} \tag{11}$$

where $C_{inv}$ is the average daily investment cost of the battery energy storage system, which can be calculated by

$$C_{inv} = (1 + \mu) \cdot (C_P \cdot P_{max} + C_E \cdot E_{max} + C_F) \tag{12}$$

where $\mu$ is the component replacement cost; $C_P, C_E$ and $C_F$ are the unit costs of power capacity, energy capacity and fixed cost, respectively.

## 3. MODEL FORMULATION

The proposed model of the BESS bidding in the pool based electricity market is described in detail. The decision variables are the capacity bids in energy market $b_{e,t}$, the capacity bids in AGC market $b_{c,t}^{up}$ and $b_{c,t}^{down}$ and the price bids in AGC market $b_{p,t}$ of the BESS for each hour in the next day.

### *3.1 Objective Function*

The bidding model is to maximise the total profit of a BESS owner, which is described as follow

$$\max \mathrm{Profit} = \sum_{t \in T} (w_e * \mathrm{Profit}_t^{\mathrm{e}} + w_{reg} * \mathrm{Profit}_t^{\mathrm{reg}} - \mathrm{Cost}_t^{\mathrm{total}}) \tag{13}$$

where $w_e$ and $w_{reg}$ are the weight of the balance between energy market and the regulation market. The weights can be predefined by the market environment and the owner of the BESS.

In the electricity market, there is a system operator between the supply companies and the retailers. The suppliers are bidding in the power pools, and the system operator makes the decision of market price and power generation offers. Since the BESSs are the price-taker in the energy market, the total revenue of a BESS in energy market $\mathrm{Profit}_t^{\mathrm{e}}$ can be calculated by (Li et al. (2011, 2017))

$$\mathrm{Profit}_t^{\mathrm{bid}} = p_t \cdot b_{e,t} \cdot \eta_{discha} \tag{14}$$

where $b_{e,t}$ is the winning power offer of the BESS at time slot $t$, termed as the capacity bidding quantity. A power supplier can only generate power if its offers are accepted. Otherwise, the extra penalties should be paid. The subscript "t" is the index of the hours in each day, since the bidding strategy is day-ahead with hourly bids in the wholesale electricity market.

In (13), $\mathrm{Profit}_t^{\mathrm{reg}}$ is the revenue of the regulation markets, which can be described as

$$\mathrm{Profit}_t^{\mathrm{reg}} = \mathrm{Profit}_t^{cap} + \mathrm{Profit}_t^{perf} \tag{15}$$

where $\mathrm{Profit}_t^{cap}$ is the revenue of the regulation capability, which can be described as

$$\mathrm{Profit}_t^{cap} = (b_{c,t}^{up} + b_{c,t}^{down}) * b_{p,t} \tag{16}$$

where $b_{c,t}^{up}$ and $b_{c,t}^{down}$ are the capacity bids; $b_{p,t}$ is the bidding price, which could influence the regulation market.

Different energy storage systems provide different regulation capacity bid. Then the system operator will make the decision and send the regulation signal to the frequency modulation unit. If the regulation bid of the BESS is accepted by the system operator, the regulation capability compensation and the regulation performance based profit can be formulated as

$$\text{Profit}_t^{perf} = \sum_{\tau=0}^{3600/\Delta T} (p_t \cdot P_\tau^{up} + p_t \cdot P_\tau^{down}) \tag{17}$$

where $\tau$ is the smaller time slot than $t$, and $\Delta T$ is the regulation period, typically from 2s - 4s. According to PJM (2011), the performance revenue is not related to the bidding capacity of the BESS, but the real-time regulation signal and the electricity price. Since each time slot, the regulation signal will only have one sign, we separate the regulation signal into regulation up signal $P_\tau^{up}$ and the regulation down signal $P_\tau^{down}$.

The total cost is calculated in (18).

$$\text{Cost}_t^{\text{total}} = C_{loss,t} + C_{ag,t} \tag{18}$$

where $C_{loss,t}$ and $C_{ag,t}$ are the charging/discharging cost and the ageing cost, respectively. Here, we doesn't considered the operation and maintenance cost since it is usually fixed.

The charging/discharging cost is the sum of the charging part and the discharging part. In this model, the charging power $P_{charge}$ is equal to the regulation down signal $P_\tau^{down}$ and $P_{discha} = P_\tau^{up}$. Therefore, the charging/discharging loss is

$$C_{loss,t} = \sum_{\tau=0}^{3600/\Delta T} p_t \cdot (P_\tau^{down}(1-\eta_{charge}) + P_\tau^{up}(\frac{1}{\eta_{discha}} - 1)) \tag{19}$$

The last part of total cost is the ageing cost, which can be calculated by (11).

### *3.2 Constraints*

*Power Constraints* In this part, the capacity limits of the BESS are considered and formulated in (20)-(22) regarding market requirements, physical constraints and regulation constraints. The sum of the capacity bids must keep within the maximum power of the BESS.

$$b_{e,t} + b_{c,t}^{up} \leq P_{max} \tag{20}$$

$$b_{e,t} - b_{c,t}^{down} \geq -P_{max} \tag{21}$$

where $P_{max}$ is the maximum output power of the BESS. It is related to the different type of the BESS.

Furthermore, the maximum regulation capacity has to be limited in a reasonable range, described in (22).

$$0 \leq b_{c,t}^{up}, b_{c,t}^{down} \leq \mu^{reg} \cdot P_{max} \tag{22}$$

where $\mu^{reg}$ is the maximum ratio of regulation capacity to the high sustained limit.

*Charging/Discharging Constraints* This part models the energy balance model of the BESS based on the physical constraints and the market requirement.

We assume that there is no energy loss during the charging/discharging process. The SOC of the BESS can be calculated as:

$$SOC_t = SOC_{t-1} + \Delta SOC \tag{23}$$

where $\Delta SOC$ is the changed SOC of the BESS between time $t-1$ and $t$. For the different type of the BESS, the charging efficiency are different. Therefore, the charging/discharging rate of the BESS ($\Delta SOC$) is expressed as

$$\Delta SOC = SOC - \eta_{charge} \cdot P \cdot h, \text{if charging} \tag{24}$$

$$\Delta SOC = SOC - \frac{1}{\eta_{discha}} \cdot P \cdot h, \text{if discharging} \tag{25}$$

where $\eta_{charge}$ and $\eta_{discha}$ are the charging/discharging efficiency of the BESS, $h$ is the time period between $t$ and $t-1$.

The BESS must keep its SOC within its energy capacity limits. According to Liu et al. (2018a), the BESS performs its best working characteristics between 20% - 80%. To get the best performance of the BESS, in this paper, the capacity limits is set as

$$\rho_{min} * E_m \leq SOC_t \leq \rho_{max} * E_m \forall t \in T \tag{26}$$

where $\rho_{min}$ and $\rho_{min}$ are the minimum and maximum efficiency operation rate. $E_m$ is the rated energy capacity of the battery storage.

*SoC Constraints* The initial and final SOC usually are set to be same during the optimization period, as described below. $t_0$ and $t_{24}$ represent the begin and end of the day.

$$SOC_{t_0} = SOC_{t_{24}} \tag{27}$$

## 4. ALGORITHM DESIGN

In this section, the proposed model of a BESS is reformulated for the reinforcement learning algorithm in detail. In the bidding market, the BESS owner needs to make the decision of bidding quantity $q_t$, bidding price $p_t$ and the capacity bids for the BESS regulation service. Hence, we set the action as

$$a_t = (b_{p,t}, b_{q,t}, b_{c,t}^{up}, b_{c,t}^{down})^T \in \mathbb{A}(s_t) = \mathbb{A} \tag{28}$$

where $\mathbb{A}(s_t)$ is the discrete action set, which is supposed to be $\mathbb{A}$ for $\forall s_t$. In the wholesale electricity market, each BESS owner only knows its own bidding quantity and price. The bidding data of the other bidders must be estimated by the previous bidding history. In this paper, the bidding quantities and prices of other rivals are presumed to be influenced by the market clearance price and the sold offer at time slot $t-1$. Some similar state-choosing methods are studied for electricity market in Li et al. (2017). Therefore, the state of the BESS owner is considered as

$$s_t = (v_{t-1}, a_{t-1}^T, SOC_t, t)^T \in \mathbb{S} \tag{29}$$

where $v_{t-1}$ is the market settlement price at time slot $t-1$. In this paper, we tend to maximise the BESS owner's profit within its bidding period, which is 24 hours of a day; therefore, time slot $t$ is set as a part of state so that the decision maker can take different actions in different hours of the day-ahead bidding strategy.

Suppose that each BESS owner has its own bidding price set and quantity set, which are defined as $\mathbb{P}_n$ and $\mathbb{Q}_n$. Therefore, the next hour state $s_{t+1} = (v_t, a_t^T, SOC_{t+1}, t+1)^T$ can be easily obtained after taking an action $a_t = a \in \mathbb{A}$. Then a reward $r_{t+1}$ is generated by state transition from $s_t$ to $s_{t+1}$, which can be considered as

$$r_{t+1} = \text{Profit}_t - C_t^W \tag{30}$$

where $r_{t+1}$ is defined under the framework of the reinforcement learning. Generally, the reward at $t+1$ time slot is the BESS owner profit in terms of the state $s_t$ and the action $a_t$ at $t$ time slot. $C_t^W$ is set as a penalty term, which is related to the local constraints, including battery and generator constraints. For example, the SOC of a BESS should be kept between 20% to 80% to obtain the efficiency operation (Rashid and Gupta (2014)). If the action leads to these inefficiency areas, the reward of this state action pair should be negative and get corresponding penalty. In this paper, the BESS owner can get the finite-time horizon reward sequence as $s_t, a_t, s_{t+1}, r_{t+1}, a_{t+1}, \cdots, s_{t+N-1}, a_{t+N-1}, s_{t+N}, r_{t+N}$, which is an episode of bidding and operation. The parameter $N$ is the trading period, which is set as 24 in this paper.

The objective of the reinforcement learning for the BESS owner $i$ is to obtain the best 24-hour reward given by

$$\mathbb{E}(\sum_{k=0}^{24} \gamma^t r_{t_0+k+1} | s(t_0) = s_0) \tag{31}$$

where $t_0$ is the initial time slot, $s_0$ is the initial state, $k$ is the time slot index, $r_i(t_0+k+1)$ is the reward based on state-action pair at time slot $t_0+k$; $\gamma$ is the discount factor which is applied to reduce the effect of future reward.

For each state-action pair, a Q function can be defined as follows:

$$Q_{t+k+1}(s,a) = Q_{t+k}(s,a) + \alpha[r + \gamma Q_{t+k}(s',a') - Q_{t+k}(s,a)] \tag{32}$$

where $a$ should be an random action under current policy $\pi$, $\alpha$ is the learning rate and $Q_{t+k}(s',a')$ is the estimate of maximum Q value related to the state-action pair $(s',a')$ at time slot $t+k$. In order to get quicker training result, the $\mathcal{E}$-greedy policy has been applied in this paper (Li et al. (2017)). To ensure the proposed algorithm can find the optimal policy which can cover maximum state values, all exploratory actions should have the probability to be chosen during the training period.

Since the states setting in the model is continuous and the dimension of the states is large, this paper applies the function approximation to solve the reinforcement learning problem. An off-policy model-free algorithm is implemented so as to find the optimal bidding strategy, which helps the BESS to get a higher profit during the trading period.

## 5. IMPLEMENTATION AND CASE STUDY

In this section, consider an electricity market with 4 BESSs, and these four BESSs bid in the AGC market to get their rewards. The planning horizon is next day 24-hour bids.

### 5.1 Datasets

In our simulation study, the real world datasets are applied to illustrate the effectiveness of our model. A 4-s based RegD & RegA signal is generated based on real RegD signal data by PJM's data set.

### 5.2 Case Implementation

In the case studies, it is supposed that all of 4 BESSs can participate in the AGC market. In this market, the BESS1

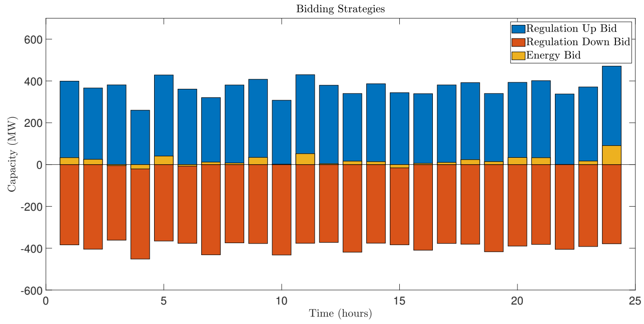


Fig. 2. The BESS optimal bids.

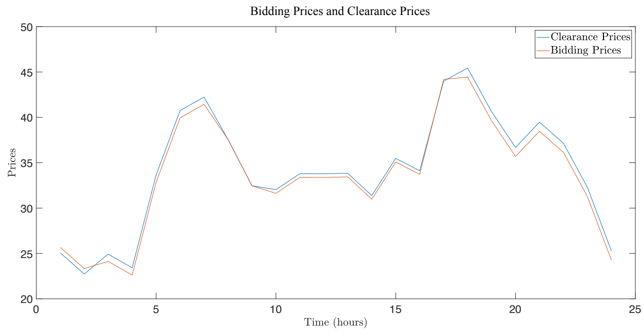


Fig. 3. The BESS optimal bidding prices.

is assumed as our own BESS, which tends to maximise the profit of next 24-hour. The decision maker of the BESS1 will implement the function approximation based reinforcement learning algorithm, seeking the proper bidding price and bidding capacities. The bidding strategies of all the other BESS are unknown and supposed to be predicted by a modelled environment of the decision maker. For the objective function shown in Eq. (22), the initial time slot is set as $t = 1$ and the end time slot is set as $t = 24$.

The relevant information and cost parameters of the BESSs are shown in Tabs. 1 - 2.

### 5.3 Results and Comparison

Figs. 2 - 3 show the optimal bidding strategies and bidding prices of the BESS in different time slots. In this case, regulation capacity dominates most of the day, since the compensation of the regulation services are high. Furthermore, we test bidding strategy of the BESS1 in regulation market and energy market with its rivals. To win the regulation services offer and earn high compensation profits, the bidding regulation price is trained to be less than the history clearance prices and the rivals' bids. When the regulation prices are very cheap, the BESS owner will purchase or sell the energy in the energy market to balance the energy loss and earn some revenue. During that period, the regulation bids are reduced because of the physical constraints of the charging/discharging rate.

The simulation results above show that the proposed model considering the ageing and transmission losses presents a more effective bidding strategy for the BESS owners in a bidding environment of multiple rivals, provides a more realistic and accurate cost-benefit result for investors as well.

Table 1. Cost coefficients for simulation studies.

| | $\mu$ | $C_P$ ($£/kW$) | $C_E$ ($£/kWh$) | $C_F$ ($£$) | $a$ ($£/MW^2$) | $b$ ($£/MW$) | $c$ ($£$) |
|---|---|---|---|---|---|---|---|
| BESS1 | 15% | 2300 | 300 | 2.58e5 | 1.25e-4 | 0.48 | 14.6 |
| BESS2 | 15% | 2250 | 450 | 2.52e5 | 1.32e-4 | 0.51 | 15.8 |
| BESS3 | 15% | 2470 | 360 | 2.49e5 | 1.28e-4 | 0.55 | 16.2 |
| BESS4 | 15% | 2320 | 280 | 2.63e5 | 1.30e-4 | 0.45 | 15.4 |

Table 2. Parameters of the BESSs for simulation studies.

| | $P_{max}$ | $E_{max}$ | $\eta_{charge}$ | $\eta_{discha}$ | $N_{100}^{fail}$ | $k_p$ |
|---|---|---|---|---|---|---|
| BESS1 | 406 | 900 | 0.868 | 0.92 | 10,000 | 0.85 |
| BESS2 | 207 | 1000 | 0.88 | 0.95 | 10,000 | 0.85 |
| BESS3 | 250 | 625 | 0.86 | 0.88 | 10,000 | 0.85 |
| BESS4 | 362 | 830 | 0.82 | 0.86 | 10,000 | 0.85 |

## 6. CONCLUSION

This paper studied the optimal bidding strategy of the BESS to maximise the profits under a multi-rival environment. We firstly proposed a bidding model for the BESS in the AGC and energy market, then solved the bidding problem with a reinforcement learning method, which uses the function approximation to avoid aggregated states and dimension curse. Simulation results verified that the proposed method can not only get a higher revenue from the AGC market, but also extend the life of the BESS and reduce the losses.

## REFERENCES

Albertus, P., Babinec, S., Litzelman, S., and Newman, A. (2018). Status and challenges in enabling the lithium metal electrode for high-energy and low-cost rechargeable batteries. *Nat. Energy*, 3(1), 16.

Coffin, D. and Horowitz, J. (2018). The supply chain for electric vehicle batteries. *J. Int'l Com. & Econ.*, 1.

Dong, Y., Zhao, T., and Ding, Z. (2019). Demand-side management using a distributed initialisation-free optimisation in a smart grid. *IET Renew. Power Gener.*, 13(9), 1533–1543.

Geaney, H. and O'Dwyer, C. (2017). Tailoring asymmetric discharge-charge rates and capacity limits to extend li-o2 battery cycle life. *ChemElectroChem*, 4(3), 628–635.

He, G., Chen, Q., Kang, C., Pinson, P., and Xia, Q. (2015). Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life. *IEEE Trans. Smart grid*, 7(5), 2359–2367.

He, G., Chen, Q., Kang, C., and Xia, Q. (2016). Optimal offering strategy for concentrating solar power plants in joint energy, reserve and regulation markets. *IEEE Trans. Sustain. Energy*, 7(3), 1245–1254.

Li, G., Shi, J., and Qu, X. (2011). Modeling methods for genco bidding strategy optimization in the liberalized electricity spot market–a state-of-the-art review. *Energy*, 36(8), 4686–4700.

Li, Z., Ding, Z., and Wang, M. (2017). Operation and bidding strategies of power plants with carbon capture. *IFAC-PapersOnLine*, 50(1), 3244–3249.

Li, Z., Ding, Z., Wang, M., and Oko, E. (2018). Model-free adaptive control for mea-based post-combustion carbon capture processes. *Fuel*, 224, 637–643.

Liu, W., Zhuang, P., Liang, H., Peng, J., and Huang, Z. (2018a). Distributed economic dispatch in microgrids based on cooperative reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.*, 29(6), 2192–2203.

Liu, Y., Xie, K., Pan, Y., Li, Y., Wang, H., Lu, W., and Zheng, C. (2018b). Lipon as a protective layer on graphite anode to extend the storage life of li-ion battery at elevated temperature. *Ionics*, 24(3), 723–734.

Michael, G., David, P., and Christopher, R. (2018). Decision on multiple-use application issues.

Monti, D., Ponrouch, A., Araujo, R.B., Barde, F., Johansson, P., and Palacín, M.R. (2019). Multivalent batteries—prospects for high energy density: Ca batteries. *Front. Chem.*, 7.

Nanduri, V. and Das, T.K. (2007). A reinforcement learning model to assess market power under auction-based energy pricing. *IEEE Trans. Power Syst.*, 22(1), 85–95.

NEE (2018). Nextera energy annual report 2018.

PJM (2011). Frequency regulation compensation in the organized wholesale power markets, dockets rm11-7-000 and ad10-11-000, u.s. fed. energy reg. comm., washington, dc, usa, oct. 2011.

Rashid, M. and Gupta, A. (2014). Mathematical model for combined effect of sei formation and gas evolution in li-ion batteries. *ECS Electrochem. Lett.*, 3(10), A95–A98.

Rashid, M. and Gupta, A. (2017). Experimental assessment and model development of cycling behavior in li-ion coin cells. *Electrochim. Acta*, 231, 171–184.

Strehle, B., Friedrich, F., Freiberg, A.T.S., Piana, M., and Gasteiger, H.A. (2019). Investigation of structural changes during long-term cycling of ncm-811 used as cathode active material in li-ion batteries. In *Meeting Abstracts*, 559–559. The Electrochemical Society.

Tan, J. and Zhang, Y. (2017). Coordinated control strategy of a battery energy storage system to support a wind power plant providing multi-timescale frequency ancillary services. *IEEE Trans. Sustain. Energy*, 8(3), 1140–1153.

Xu, B., Oudalov, A., Poland, J., Ulbig, A., and Andersson, G. (2014a). Bess control strategies for participating in grid frequency regulation. *IFAC Proc. Vol.*, 47(3), 4024–4029.

Xu, Y., Zhang, W., Hug, G., Kar, S., and Li, Z. (2014b). Cooperative control of distributed energy storage systems in a microgrid. *IEEE Trans. Smart Grid*, 6(1), 238–248.

Ying, W., Zhi, Z., Botterud, A., Zhang, K., and Qia, D. (2016). Stochastic coordinated operation of wind and battery energy storage system considering battery degradation. *J. Mod. Power Syst. Clean Energy*, 4(4), 581–592.

Zhou, X., Stein, J.L., and Ersal, T. (2017). Battery state of health monitoring by estimation of the number of cyclable li-ions. *Control Eng. Practice*, 66, 51–63.