Continuous Intraday Trading: An Open-Source Multi-Market Bidding Framework for Energy Storage Systems

Kim K. Miskiw
Karlsruhe Institute of Technology
Karlsruhe, Germany
kim.miskiw@kit.edu

Jan Ludwig
Karlsruhe Institute of Technology
Karlsruhe, Germany
jan.ludwig@student.kit.edu

Leo Semmelmann
Karlsruhe Institute of Technology
Karlsruhe, Germany
leo.semmelmann@kit.edu

Christof Weinhardt
Karlsruhe Institute of Technology
Karlsruhe, Germany
weinhardt@kit.edu

Abstract

The increasing integration of renewable energy sources and the growing need for flexibility have made trading opportunities close to delivery increasingly important in European energy markets. This shift creates new profit opportunities for market participants, such as energy storage systems, which capitalise on temporal price differences. However, it also necessitates effective coordination across multiple markets—a significant challenge, particularly with continuous intraday trading. To address this, we propose an open-source, implementable framework for multi-market bidding under uncertainty designed to increase the profitability of energy storage systems through enhanced coordination. Specifically, we consider two spot markets: the day-ahead market and continuous intraday trading. By exploring varying degrees of market coordination with mixed integer linear optimisation, a rolling intrinsic algorithm and deep reinforcement learning, we discuss its ability to capture the benefits of multi-market participation. Already, a myopic multi-market approach that combines the day-ahead market and continuous intraday trading shows decreased risk at similar profit levels. The anticipation of rolling intraday profits, called coordinated multi-market bidding, remains a challenging task, but we present cases in which it starts to become beneficial. Most notably, we provide an open-source framework to further analyse the market dynamics.

CCS Concepts

• **Computing methodologies → modeling and simulation, Machine learning**.

Keywords

Battery Storage, Multi-Market Bidding, Day-Ahead Market, Continuous Intraday Market

ACM Reference Format:
Kim K. Miskiw, Jan Ludwig, Leo Semmelmann, and Christof Weinhardt. 2025. Continuous Intraday Trading: An Open-Source Multi-Market Bidding Framework for Energy Storage Systems. In *The 16th ACM International Conference on Future and Sustainable Energy Systems (E-ENERGY '25), June 17–20, 2025, Rotterdam, Netherlands.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3679240.3734631

E-ENERGY '25, Rotterdam, Netherlands

ACM ISBN 979-8-4007-1125-1/25/06
https://doi.org/10.1145/3679240.3734631

1 Introduction

For the transition to a renewable energy system, large-scale storage systems are essential: the intermittency of renewable energy sources necessitates storage to maintain the balance between supply and demand [8]. Lithium-ion batteries are widely regarded as a promising solution for renewable energy integration, offering high power and energy density along with a high efficiency [26, 29]. However, the deployment of Battery Energy Storage System (BESS) capacity in many advanced power markets is progressing at a modest pace. For instance, in Germany, the installed BESS capacity reached approximately 12 GWh by the end of 2024, while projections for a renewable energy-based system estimate the need for as much as 111 GWh [11].

Recent developments suggest a shift in this trajectory. Transmission system operators in Germany have received BESS connection requests totalling 161 GW for the coming years [28]. This surge in interest highlights the growing recognition of the economic potential of BESS. The key revenue stream for BESS is temporal arbitrage—buying electricity during low-price periods and selling it during high-price periods [4]. Temporal arbitrage can be executed across different electricity markets, i.e., the Day-Ahead (DA) market or the Intraday Continuous (IDC) market in the European energy market context [38]. Most existing studies in the field focus exclusively on bidding strategies for either the DA auction [9, 10, 39, 59] or the IDC market, which operates with an open order book for matching buyers and sellers [2, 27]. In practice, however, BESS owners can participate in both markets, first submitting bids to the DA auction and then adjusting their positions in the IDC market based on real-time conditions [38].

Studies addressing this multi-market bidding strategy remain limited. Many either overlook the IDC market [31, 33, 46], simplify it to a single-auction framework [34] or a small number of auctions [51]. This simplification is problematic because the IDC market—and specifically the ability to reposition financially throughout the day—often provides higher profit potential than the DA market alone [52]. Other studies that model trading on both the DA auction and continuous trading on the IDC rely on multistage stochastic problems [38], which provide general insights to multi-market behaviour but are difficult to transfer to the real-world decision problems of practitioners.

However, modelling realistic trading operations of BESS within the multi-market bidding problem scope is critical for policymakers and practitioners. This is especially relevant given the recent surge in connection requests for large-scale BESS. First, such a realistic multi-market modelling process provides a more accurate quantification of the profitability potential of BESS, supporting informed investment decisions and potentially accelerating the deployment of BESS capacity. Second, they assist BESS trading firms in optimising their decision-making processes across multiple markets, reflecting real-world complexities and trading opportunities. Third, these models enable policymakers and researchers to assess the impact of regulatory changes on market participation and trading volumes, facilitating better-informed policy design to support the integration of BESS into energy markets.

Our study introduces a multi-market bidding framework for large-scale BESS designed to model real-world trading processes under uncertainty and realistic conditions. The framework enables a consideration of the multi-market optimisation problem in two ways. First, as a myopic bidding decision (first DA market auction, then participating with the buy and sell positions on the IDC market), for which we extend the work in [52] to consider transaction costs and the DA market. Second, we enable the learning of a coordinated decision (participating in the DA market with anticipation of profit potentials on the IDC market). For the myopic bidding, the DA trading decisions are optimised using a Mixed-Integer Linear Programming (MILP), while the IDC market is represented by a Rolling Intrinsic Trading Strategy (RI) algorithm that allows for ongoing repositioning throughout the trading day. The coordinated approach is modelled with a Markov Decision Process (MDP) incorporating the RI algorithm, which is then solved using Deep Reinforcement Learning (DRL). The framework incorporates uncertainties in price realisations, auction outcomes, and price developments on the IDC and the DA markets, providing increased modelling accuracy and realism. In addition to the comparison of myopic and coordinated multi-market bidding, we benchmark trading outcomes and profitability against single-market participation on either the DA or IDC market.

In this context, our study makes the following contributions. We present a comparison between the market performance of single- and multi-market spot bidding strategies with BESS. We show that while solely bidding on the IDC yields higher profits during the investigated period, multi-market bidding profits have the potential to mitigate risks of lower daily returns. Furthermore, we investigate differences in profit between myopic and coordinated approaches for multi-market bidding. While in our implementation the myopic approach outperforms the coordinated one, we show the potential of the latter for superior performance in particular cases. We publish our framework open-source, to enable further research on BESS spot market bidding strategies.

Our study is structured as follows: Section 2 reviews the existing literature on single- and multi-market bidding strategies for BESS. Based on this review, we detail our proposed methodology in Section 3. In Section 4, we present the case study involving a large-scale lithium-ion BESS in Germany, incorporating realistic technical constraints and transaction costs. Section 5 presents the results, which are further analysed and discussed in Section 6. Finally, we conclude our study in Section 7.

2 Literature

This paper addresses two closely related short-term operational challenges for generation and storage capacities: the optimal dispatch or unit commitment problem for a BESS, and the optimal trading or bidding problem. While these issues are distinct, they are often interconnected. The unit commitment problem focuses on adhering to a predetermined schedule or providing flexibility on demand at the lowest possible cost, primarily addressing technical constraints. In contrast, the bidding problem incorporates (expected) market outcomes, optimising bids to enable profit-maximising operations [34]. The objective of this paper is to maximise profits without predetermined market commitments and, hence, the unit commitment problem is treated implicitly. Instead, the focus lies on the bidding problem with integrating technical constraints. Accordingly, the following literature review concentrates on studies addressing this focus. We provide an overview tabel of the most relevant studies in Appendix A.1.

Recent studies have examined single revenue streams within liberalised electricity markets in isolation with a variety of methods [31, 38]. We provide an overview of studies focusing on single revenue streams in the following.

Day-ahead: To optimise revenue in single-market operations, recent research primarily targets exploiting temporal arbitrage opportunities, especially through participation in the DA auction, where energy is traded the day before delivery. Most studies in this area focus on the formulation of a profit optimisation problem [9, 10, 39, 59]. These studies commonly rely on a DA price forecast, upon which charging and discharging (and therewith connected bidding) decisions of the BESS are optimised. Other studies use a DRL-based approach to derive bidding [7, 24, 60, 61, 63]. Unlike traditional optimisation methods, DRL is model-free, dynamically learning to handle uncertainty through interactions with the environment [37, 40].

Intraday: In contrast to the DA market, which is used for scheduling electricity generation and consumption on an hourly basis a day in advance, intraday markets focus on short-term adjustments to balance supply and demand closer to real-time operations [57]. This is especially becoming more important due to the increasing share of intermittent renewable generation. IDC trading is characterised by high volatility, which offers significant opportunities for exploitation by BESS. Current literature often simplifies revenue estimation by using descriptive indices such as ID1, ID3, or IDFull [17], which represent the weighted average prices of all continuous trades during the last hour, the last three hours, or the entire trading period for an IDC product, respectively. They use those, together with a MILP, to estimate the arbitrage potential within the IDC market [6, 30, 49]. Other literature uses dynamic programming to derive bidding in the intraday market. Yet they simplify the continuity of the market [27], as they do not account for products with different times to maturity. These abstractions fail to fully capture the true potential of the price fluctuations that occur during the trading periods of these products. The initial version of the RI is introduced by [23]. The authors in [52] recently extended the algorithm. The algorithm employs a RI trading model, leveraging discretised transaction data to iteratively optimise trading decisions within defined time buckets, ensuring adherence to BESS constraints while capitalising on continuous market price fluctuations. This approach enables dynamic re-optimisation, allowing the system to adapt to real-time price changes and exploit arbitrage opportunities that arise over time. The authors conduct a case study demonstrating the effectiveness of the algorithm by comparing its performance with index-based methods, demonstrating their underestimation of possible BESS profits.

Multi-market: Focusing on single revenue streams in isolation is a vast simplification of electricity market trading as highlighted by [38] and [31]. Consequently, there is an increasing interest in multi-market bidding strategies in literature. Developing effective multi-market strategies for BESS remains challenging, for example, due to the high volatility of market prices and the difficulty of allocating BESS capacity to maximise revenue without exceeding capacity limits [37]. Nevertheless, multi-market participation has proven beneficial for generation portfolios, including BESS, as demonstrated in studies such as [21, 31, 33, 38, 41, 46, 57]. While existing contributions highlight the benefits of multi-market bidding, there remains a lack of algorithms for joint-market participation under price uncertainties for BESS [37]. Current studies often neglect the IDC entirely [31, 33, 46], simplify it to a single auction [21, 25, 34, 41], or limit their analysis to a few discrete auctions [51, 53, 57]. Most approaches use optimisation-based methods under uncertainty, which are computationally demanding and face scalability challenges due to the curse of dimensionality as considered markets increase [37, 40]. Although some studies report limited potential for coordinating multiple markets, they also fail to incorporate the continuous nature of the IDC, underscoring the importance of addressing this gap [43]. With stochastic programming, continuous trading on the IDC, combined with the DA market, across 34 decision stages is incorporated [38]. Their results remain primarily theoretical, relying on simulated price distributions from 2015 to 2018 and focusing their analysis on one summer and one winter day in 2017. Their derived bidding strategies inform the design of our approach.

As an alternative to optimisation-based approaches, DRL has emerged as a promising method for addressing multi-market decision-making challenges under uncertainty [37, 40]. While most studies focus on single-market settings [36], a growing body of research explores multi-market applications [3, 12–14, 37, 54]. However, as summarised in the table in Appendix A.1, these studies often rely on outdated market data [3, 13, 37] and do not model the IDC [3, 13, 14, 37]. Many multi-market bidding applications assume full observability, incorporating uncertain market prices with perfect foresight in the state space [3, 45]. While renewable generation uncertainty is more frequently addressed [36], state spaces typically include exogenous demand or renewable generation [12, 13, 36]. Some studies also rely on limited, discretised action spaces [12, 14], often justified by experiments showing that increased action space complexity reduces algorithm performance [63]. However, stochastic programs for multi-market bidding with the IDC reveal that anticipating IDC profits encourages spreading charging and discharging across multiple hours for greater nuance [38]. The only study considering the IDC is [12]. They use a heuristic to determine preceding markets such as the DA. Therefore, the interrelation between the markets is not subject to learned behaviours.

In line with this research gap, we propose a multi-market bidding algorithm based on DRL that integrates the IDC and utilises finer discretization of the action space, drawing on insights from multi-market bidding with stochastic programming. Our approach combines the industry-standard RI IDC algorithm [48] with a DRL method both under full uncertainty, leveraging publicly available price forecasts. Additionally, we apply this algorithm to recent market data from 2019 to 2023.

3 Methodology

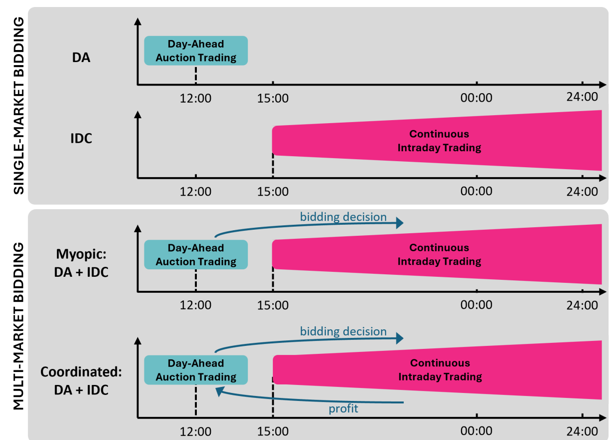


Figure 1: Schematic representation of different single- and multi-market modelling approaches.

In the European electricity market, electricity is traded on power exchanges or over-the-counter. This study focuses on spot markets within a power exchange since the goal is to derive a revenue optimisation strategy for BESS, allowing the exploitation of short-term price volatility. There are three spot market auctions in European countries, from which we present Germany's specific opening and closing times. Two of them will be analysed in detail in this study, namely the participation in the DA auction and in IDC trading. To derive a trading strategy for a BESS across both selected spot markets, we present the modelling approach step by step. Firstly, we model that the BESS is solely bidding on one spot market, namely first the DA market (Section 3.1) and then the IDC market (Section 3.2). We then move on to model bidding on both considered markets (Section 3.3). We present two multi-market bidding strategies with a different degree of complexity. The myopic strategy considers a subsequent decision: first, bidding decisions for DA are made. Then, the conducted trades serve as input for the IDC bidding problem. Whereas, the coordinated approach already models profit expectations of the IDC during the DA bidding process. The different approaches are visualised in Figure 1.

3.1 Day-ahead Auction Bidding

The DA auction allows market participants to trade electricity for each hour of the next day. Market participants can submit quantity-price bids for different hours until 12:00 p.m. CET the day before delivery. After the submission window closes, market clearing takes place. The pricing model used is pay-as-cleared [15]. For the DA market, we present a forecast-based MILP model and a DRL approach where we model the BESS in the DA market as a MDP. For similar MDP formulations, see sources in Appendix A.1.

3.1.1 Mixed Integer Linear Program. The general bidding decision of a BESS on the DA can be formulated as the following optimisation problem. On the DA market, the BESS maximises their profit under the expectation of a forecasted electricity price $\widehat{p}_t$ for the hours ($t \in T$) of the upcoming day, as seen in objective function (1). The BESS chooses the amount it sells q_t^{sell} (discharges) and buys q_t^{buy} (charges) in every time step t.

$$\text{maximise} \quad \sum_{t \in T} \left(\widehat{p}_t \, q_t^{\text{buy}} - \widehat{p}_t \, q_t^{\text{sell}} \right) \tag{1}$$

subject to:

$$q_t^{\text{charge}} = \eta \, q_t^{\text{buy}} \quad \forall t \in T \tag{2}$$

$$q_t^{\text{discharge}} = \frac{q_t^{\text{sell}}}{\eta} \quad \forall t \in T \tag{3}$$

$$\text{SoC}_t = \text{SoC}_{t-1} + \frac{q_{t-1}^{\text{charge}}}{C^{\text{bess}}} - \frac{q_{t-1}^{\text{discharge}}}{C^{\text{bess}}} \quad \forall t \in T, \ t > 0 \tag{4}$$

$$\text{SoC}_0 = 0 \tag{5}$$

$$\sum_{t \in T} \frac{q_t^{\text{charge}}}{C^{\text{bess}}} - \sum_{t \in T} \frac{q_t^{\text{discharge}}}{C^{\text{bess}}} = 0 \tag{6}$$

$$q_t^{\text{buy}} \leq \alpha_t^{\text{buy}} \, P^{\max} \quad \forall t \in T \tag{7}$$

$$q_t^{\text{sell}} \leq (1 - \alpha_t^{\text{buy}}) \, P^{\max} \quad \forall t \in T \tag{8}$$

$$0 \leq \text{SoC}_t \leq 1 \quad \forall t \in T \tag{9}$$

$$\sum_{t \in T} \left(q_t^{\text{charge}} + q_t^{\text{discharge}} \right) \leq 2 \, C^{\text{bess}} \tag{10}$$

$$\alpha_t^{\text{buy}} = \begin{cases} 1 & \text{if } q_t^{\text{buy}} > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

The model is subject to several technical constraints that govern the BESS operations. The total energy capacity of the BESS, denoted as C^{bess}, represents the maximum amount of energy that can be stored, measured in MWh. The maximum power that the BESS can charge or discharge is represented by $P^{\max}$, measured in MW. The constraints (2) and (3) define the charging and discharging quantities at each time step, taking into account the BESS efficiency η. The State of Charge (SoC) refers to the energy stored in the BESS, where $0 \leq \text{SoC} \leq 1$, with zero indicating empty, and one indicating fully charged. The effect of charging and discharging on the SoC is modelled with Equation (4), ensuring that the initial SoC (5) and final SoC (6) conditions are met. Additional constraints ensure that there is no simultaneous buying and selling (7), (8). Therefore, a binary indicator α^{buy} is introduced (11). In addition, these constraints ensure alignment with the maximum charging and discharging capacity $P^{\max}$, which we assume is equal for charging and discharging, consistent with the assumptions made in [63]. Equation (9) ensures that the SoC stays within operational limits. Finally, the cycle limit constraint (10) ensures that cycling constraints are met by limiting the total energy throughput. A full cycle is defined as one full charge and one full discharge. However, a full cycle does not have to occur continuously; it can accumulate over several partial charge and discharge events.

This formulation identifies the volume bids (q_t^{sell}, q_t^{buy}) that optimise the profit given a forecasted electricity price. Note that we assume that all of our bids have been accepted. This could be made possible by bidding on the price limits of the auction. Namely -500 € when selling and 4000 € when buying [18]. This is a simplification often made in literature[43] and has limited effect since the auction is settled with a pay-as-cleared mechanism (cf. Table 1 in Appendix A.1).

3.1.2 Markov Decision Process. In the following section, we formalise the foregoing optimisation problem as a MDP and specify how it can be solved with DRL. The chosen representation is based on the approaches presented in related literature [7, 24, 35, 56, 58, 60, 61, 63]. This section introduces all design choices for this MDP, which will be further developed to consider multiple markets. A MDP is defined by the tuple (S, A, P, R, γ), where S is the state space representing both the operational state of the BESS and the market conditions that can be observed by the DRL agent. In our case, the DRL agent is the operator of the BESS. A is the action space corresponding to the set of possible actions available to the agent. The state transition probabilities are given by $P(s_{t+1} \mid s_t, a_t)$, while $R(s_t, a_t)$ denotes the reward function, which represents the agent's reward at each time step t. The discount factor γ captures the degree to which future rewards are valued.

To guide the chosen action by the agent, we use the following state space S at each time step t:

$$s_t = \left[q_{t-1}, \text{SoC}_t, \text{rc}_t, \sin\left(2\pi \frac{t}{24}\right), \cos\left(2\pi \frac{t}{24}\right), \widehat{p}_t, \widehat{p}^T \right]. \tag{12}$$

The sold or bought quantity of the last step q_{t-1} is included in the state space.[1] The SoC of the BESS at the time step t is denoted as SoC_t. The cycle restriction (10) is indicated in the state space by including the remaining cycles of the BESS at time t, rc_t. The terms $\sin\left(2\pi \frac{t}{24}\right)$ and $\cos\left(2\pi \frac{t}{24}\right)$ are used to encode the time step of the day, capturing periodic patterns over a 24-hour cycle. $\widehat{p}^T$ represents the transposed scaled forecast electricity prices for the entire day, while $\widehat{p}_t$ denotes the scaled forecast electricity price specifically at time step t. The forecast price vectors are scaled using min-max scaling.

The **action space** A is discrete and is represented as

$$a_t \in \{a^{\text{charge}}, a^{\text{idle}}, a^{\text{discharge}}\}, \tag{13}$$

where a_t denotes the operational decision for the BESS at time step t. The action a^{charge} represents charging the BESS with $P^{\max}$ (implicitly reflecting (7)). Meanwhile, a^{idle} indicates that the BESS remains in idle state. The action $a^{\text{discharge}}$ indicates that the BESS is being discharged with $P^{\min}$ (implicitly reflecting (8)).

In the MILP (6) and (10) limit the behaviour of the SoC to only include one full cycle. This is handled in the MDP by the remaining

[1] Please note, similar to the presented MILP, the bid price is not subject to a decision here but is set to the minimum or maximum price. Hence, all bids are accepted, and including this information means there is no information leakage. The MDP still acts under full uncertainty.

cycles $\mathrm{rc}_t \in [0,1]$, which represent the fraction of the BESS usable cycles that have been consumed over time. They depend on the allowed cycles per period. To ensure better convergence of the model, if the chosen actions violate technical constraints of the SoC, compare (9), or the cycle restriction, compare (10), they are changed to the closest technically feasible action values q_t. This is also called action clipping and is in line with the approaches presented in related literature [24]. The quantity sold or bought from the market $q_t \in [-1,1]$ is therefore similar as in the optimisation problem $(q_t^{\mathrm{sell}}, q_t^{\mathrm{buy}})$ expressed by (14).

$$q_t = \begin{cases} -\min\left(P^{\max}, C^{\mathrm{bess}} - (\mathrm{SoC}_t\, C^{\mathrm{bess}}), \mathrm{rc}_t\right) & \text{if } a_t = a^{\mathrm{charge}} \\ \min\left(P^{\max}, (\mathrm{SoC}_t\, C^{\mathrm{bess}}), \mathrm{rc}_t\right) & \text{if } a_t = a^{\mathrm{discharge}} \\ 0 & \text{if } a_t = a^{\mathrm{idle}}. \end{cases} \tag{14}$$

Based on the submitted trade quantities, the agent receives a reward. The **reward function** $R(s_t, a_t) \in [-1,1]$ at each time step is derived from the scaled revenues from trading in the DA auction, calculated as the product of the realised quantity and the scaled realised market price. Therefore, we use scaled realised prices p_t as in [7, 60, 63], in particular min-max scaling.[2]

The cumulative reward over all time steps R is then given by:

$$R = \sum_{t=0}^{T} \gamma\, p_t\, q_t \tag{15}$$

where $\gamma \in [0,1]$ is the discount factor, p_t it the realised market price and T is the episode length (24 hours).

The BESS dynamics and market conditions determine the transitions between states $P(s_{t+1} \mid s_t, a_t)$. In our case, the transitions are deterministic. That is, the next state s_{t+1} is uniquely determined by each state-action pair. After an action is performed at time step t, the BESS's SoC is updated based on the action, and the time step advances similarly as it is shown in (4). When one cycle is finished, the episode is terminated, as more trading would violate the cycling constraint.

3.2 Continuous Intraday Bidding

Afterwards, the IDC trading begins at 4:00 p.m. CET on the day before delivery. Orders can be placed from 4:00 p.m. CET until the gate closes for each product. For example, the quarter-hourly product with delivery at 02:00 a.m. CET on the day of delivery remains tradable until 01:55 a.m. CET, just five minutes before the delivery period [15]. This market structure allows for real-time adjustments of a participant's position, allowing for the BESS positions to respond to volatility in renewable energy forecasts or to take advantage of arbitrage opportunities [57]. For an order to be executed, another participant must match it in price and volume. There are three types of products available in the IDC market, namely power delivery for hourly, half-hourly and quarter-hourly time intervals of the next day. In this paper, we will focus on quarter-hourly products. We choose this product as it offers the highest profit for the RI participation [52] and its consistently growing importance and liquidity since 2019 [50].

For modelling the IDC market, we employ the RI algorithm from [52], an industry-standard approach that exploits lucrative price spreads in a completely risk-free manner [48]. This risk handling contrasts the approach in [12], where open positions can be made in the IDC speculating on future prices. The RI algorithm systematically iterates over predefined time buckets that discretise the trading time horizon and continuously optimise trading decisions based on the available price data. The goal is to exploit intraday price fluctuations while considering the BESS system's technical limitations. In the following, we summarise the data transformation and the actual trading algorithm. For a detailed description of the open-source algorithm, we refer to [52].

The continuous nature of the IDC market and the irregular timing of trades make a detailed analysis of the full order book very complex, and it is computationally expensive [52]. To overcome this, the algorithm uses time discretization into a finite number of buckets, while for each, the Volume Weighted Average Prices (VWAPs) are used as a price. For example, if the bucket size is set to 60 min. The trading period, which spans from 16:00 p.m. CET on the day before delivery to 23:55 p.m. CET on the day of delivery, is divided into 32 buckets (bucket 1: 16:00-17:00 p.m. CET, bucket 2: 17:00-18:00 p.m. CET, ..., bucket 32: 23:00-00:00 a.m. CET) during which each product can be traded. We further denote the available buckets as $k \in K$, where $K \in \{1,2,3,\ldots,32\}$ in the foregoing example.

Algorithm 1 Rolling Intrinsic Algorithm

1: **Initialisation**: Battery and trading parameters are initialised. These parameter sets are provided as inputs to the model.
2: **for all** k **in** K **do** ▷ Iterate through all buckets of the time period
3: Retrieve the set $\mathrm{VWAP}_k^{\mathrm{valid}}$ of tradable products in the given time period k and threshold θ.
4: Identify the most profitable combination of potential trades in period k, given tradable products $\mathrm{VWAP}_k^{\mathrm{valid}}$, using an optimisation model.
5: Append these trades to the list of previous trades, serving as inputs for subsequent periods.
6: **end for**
7: **Output**: The set of trades is returned and can be used to calculate metrics such as full load cycles or gross revenues generated by trading.

This discretised data is now used to iterate over the entire trading horizon, bucket-by-bucket, and apply a MILP optimisation. The optimisation is performed based on all tradable products and their price estimated by the VWAPs in one bucket. It optimises the revenue of the BESS by taking advantage of lucrative repositioning opportunities without relying on any assumptions about future

[2]The minimum price is thus scaled to zero, and the maximum price is scaled to one. Assuming a C-rate of one for this example, meaning that the BESS can be fully charged and discharged within one hour, the maximum cumulative reward over each day is one, since buying at the cheapest price results in a reward of zero and selling for the highest price results in a reward of one. This means that if the agent's decisions only realise a spread of 80 € on a day when the maximum spread is 100 €, the reward would be 0.8. If the agent makes the worst decision, buying at the highest price and selling at the lowest price, the reward would be minus one.

price movements, as it only focuses on the current bucket. It is assumed that if the algorithm places an order at the calculated VWAP price, a trade will be matched by a counterpart and, hence, executed. As this is a strong assumption, the authors introduce a threshold θ of at least 10 trades that must be available in a bucket k for a valid VWAP ($\mathrm{VWAP}_k^{\mathrm{valid}}$) to be calculated. If there are not enough trades, the lack of liquidity is represented by a missing value. We quote the exact representation of the algorithm's pseudocode in Algorithm 1 from the paper [52].

3.3 Multi-Market Bidding

In this section, we focus on multi-market bidding, for which the foregoing implementations serve as a basis. The DA market, characterised by higher trading volumes and liquidity in Germany, offers predictable prices with lower sensitivity to price impacts [41, 62]. In contrast, the IDC market provides greater profit potential due to its high price volatility but involves increased complexity. Price movements in the IDC are driven by rapidly changing conditions, such as renewable energy forecast updates, making reliable predictions challenging [42]. To address these complexities, we first present a myopic strategy that integrates both markets independently. This serves as a foundation for the subsequent coordinated approach, which aims to anticipate future revenues of the IDC when making decisions on the DA market and optimise bidding decisions across markets.

3.3.1 Myopic. Myopic bidding, a term commonly used in stochastic programming, refers to a multi-market bidding strategy with limited foresight, focusing solely on immediate profits [38]. In our market setup, this approach involves first optimising bids for the DA market based on forecasts without accounting for potential future profits in the IDC market. Subsequently, the IDC market is utilised to enhance revenues by exploiting additional arbitrage opportunities while adhering to existing DA positions. The original RI algorithm has to be adapted in the following way to enable this stacking behaviour. During the initialisation in Algorithm 1 the DA trades from the MILP optimisation in Section 3.1 are included. Note that the RI algorithm can perform counter trades against the original DA trades, which exclude them in the resulting dispatch pattern. We present an example day of the results of the RI with DA trades in Appendix A.2 to visualise its concepts.

While this method enables an easy form of multi-market bidding under full uncertainty, it misses profit opportunities, as highlighted by [6]. These missed opportunities include inter-market arbitrage and the strategic avoidance of DA positions when extreme and potentially more profitable IDC price movements are anticipated. DA positions can restrict the flexibility of the applied RI algorithms, potentially reducing overall profitability across both markets.

3.3.2 Coordinated. These shortcomings are addressed in the second multi-market bidding strategy. It does so by enabling the anticipation of the profit potential in the IDC when making decisions for the DA market. By striking a balance between the two markets, this approach should not only increase profit generation but also help hedge risks by diversifying across both markets, allowing for a more robust strategy than forecast-based MILP approaches. To enable such an adaptation, we further develop the presented DRL-based approach in Section 3.1 and combine it with the RI algorithm. While the RI handles the trading on the IDC, we include the impact of these trades in the reward of the DRL agent. This way, the reward is not only determined by the profit in the DA auction, but also by the IDC. For example, if no DA trades are made, the agent will still receive a considerable positive reward if profits from the RI arise, which could incentivise waiting for the IDC. The final reward at each step t is now calculated as the sum of the DA and RI rewards. Both individual rewards are scaled with the average profit of the single-market bidding on the IDC.[3]

To support the learning process, we introduce more features that potentially signal the variance of the IDC revenue potentials in the observation space as summarised in (16). We loosely base these features on the foregoing correlation analysis of the daily IDC profits with potential features. New features in this state space are the month $m : D \to \{1, 2, \dots, 12\}$ such that $m(d)$ maps each day $d \in D$ to its corresponding month. The day of the week, $wd : D \to \{1, 2, \dots, 7\}$, maps each day d to an integer representing the day of the week, where 1 is Monday and 7 is Sunday. For each holiday d_h, we set $wd(d_h) = 7$, which encodes holidays as Sundays. All these time variables are again encoded by the sine and cosine. In addition, we add the daily mean, standard deviation, minimum and maximum of the 24-hour lagged spread between DA prices and the IDFull s_{d-1}. This price index (IDFull) is the weighted average price of all continuous trades executed during the full trading time [17]. This feature must be lagged by 24 hours because we operate under uncertainty, not using intraday information since it is not known during the inference of the model. Finally, we include information about the daily variability of the renewable generation and load to enhance the anticipation of expected price movements on the IDC. For this, we include the daily mean $\bar{\hat{g}}_d^{\mathrm{wind}}$ and standard deviation $\sigma(\hat{g}_d^{\mathrm{wind}})$ of the forecasts for aggregated onshore and offshore wind infeed $\hat{g}_t^{\mathrm{wind}}$. The photovoltaic infeed $\hat{g}_t^{\mathrm{pv}}$ and the load $\hat{l}_t$ are not explicitly included due to the correlation identified in a preliminary analysis. Furthermore, we include hourly values that indicate the change in forecasted load $\Delta\hat{l}_t$, photovoltaic $\Delta\hat{g}_t^{\mathrm{pv}}$, and wind infeed $\Delta\hat{g}_t^{\mathrm{wind}}$ compared to the preceding hour, reflecting the current momentum of the price-influencing factors.

$$
\begin{aligned}
s_t = \big[\, & q_{t-1},\ \mathrm{SoC}_t,\ \mathrm{rc}_t,\ \widehat{p}_{d,t},\ \widehat{p}_d^T,\ \sin\!\Big(2\pi \tfrac{t}{24}\Big),\ \cos\!\Big(2\pi \tfrac{t}{24}\Big), \\
& \sin\!\Big(2\pi \tfrac{m(d)}{12}\Big),\ \cos\!\Big(2\pi \tfrac{m(d)}{12}\Big),\ \ wd(d), \\
& \sigma_{s_{d-1}},\ \bar{s}_{d-1},\ s_{d-1}^{\min},\ s_{d-1}^{\max}, \\
& \bar{\hat{g}}_d^{\mathrm{wind}},\ \sigma(\hat{g}_d^{\mathrm{wind}}),\ \Delta\hat{l}_t,\ \Delta\hat{g}_t^{\mathrm{wind}},\ \Delta\hat{g}_t^{\mathrm{pv}}\ \big].
\end{aligned} \tag{16}
$$

We extend the action space to enable more discrete actions. The introduced DA implementation and most studies rely on discretised action spaces with three distinct values [12, 14], which reflects charging/discharging at full capacity or idling. For single-market scenarios, experiments show that increased action space complexity (e.g., 3, 5, 9, or 15 discrete steps) even reduces made profits [63].

[3]To accurately account for the clearing times of both markets, the final combined reward is calculated only after completing 24 DA steps. This reward is then stored in the experience, ensuring no information leakage occurs and all decisions are made under full uncertainty.

However, as stochastic programs for multi-market bidding with the IDC reveal that anticipating profits in the IDC encourage a shift from trading full charging/discharging capacity in one hour to spreading this across multiple hours [38]. To enable a similar learned behaviour, we increase the freedom of the modelled DRL approach by increasing the action space (17). The choice of discrete charging/discharging steps aims to strike a balance between the finding that increasing the number of discrete steps reduces training performance and the fact that coordinated bidding encourages a shift from training at full capacity all at once. The intervals are chosen to be equidistant, as we found no evidence to support particular other distances.

$$a_t \in \{-P^{\max}, -\frac{2}{3}P^{\max}, -\frac{1}{3}P^{\max}, 0, \frac{1}{3}P^{\max}, \frac{2}{3}P^{\max}, P^{\max}\}. \tag{17}$$

Further, the trading task for the DRL implementation of the coordinated multi-market bidding is increased gradually. This means we first train only on the DA market to learn general charging and discharging patterns, and then incorporate the IDC profit and continue the training. This is generally referred to as curriculum learning. The concept is similar to transfer learning, where knowledge is transferred from one task to another. When the sequence of tasks has an increasing complexity, it can be defined as a curriculum. Learning based on such a curriculum has been shown to improve performance on a complex problem as well as reduce convergence times [44].

4 Case Study

This section details the design and implementation of a case study to validate the applicability of the approaches presented in Section 3. To this end, we model a BESS with technical specifications defined in Appendix A.3 and apply it to the German DA auction and IDC market. In contrast to other studies [38, 54], we account for IDC transaction costs to prevent the RI algorithm from exploiting any spread above zero. The costs are based on the EPEX SPOT fee schedule for November 2024, comprising 0.09 €/MWh per trade and an additional 0.01 €/MWh for clearing costs [16].

Our proposed multi-market bidding model requires three types of data: forecast data, realisations of auction prices, and IDC transaction data. Forecasts include German wind generation (ENTSO-E) [19] and Energy Exchange Austria (EXAA) prices [20], used here as a proxy for German DA auction prices to avoid the need for synthetic forecasting models. In [62], it is demonstrated that early price information from the Austrian Energy Exchange, which settles prior to the German DA, serves as a reliable proxy for forecasting the results of the German auction. Auction prices and IDC transaction data are sourced from ENTSO-E and EPEX SPOT SE, respectively [16, 19]. Transaction data, which captures only executed orders, are preferred over full order book data due to computational constraints, as justified in [38, 41, 52]. The dataset covers January 1, 2019, to December 31, 2023, but results here focus on November and December 2020.

All calculations are performed on a virtual machine with 16 VCPUs and 32 GB RAM. Optimisation problems are solved with PuLP and the open-source solver CBC [22]. The DRL model for coordinated multi-market bidding was trained using Proximal Policy Optimisation (PPO) from Stable Baselines3 [47], chosen for its robustness in handling both discrete and continuous action spaces. Details of the PPO parameterisation are listed in Table 2, and the complete source code is available open-source under https://github.com/kim-mskw/BessBidder.git. We use 10% of the model data as a test data set. After ruling out overfitting by comparing the performance on the test and train data set, we present the results for the entire model horizon.

5 Results

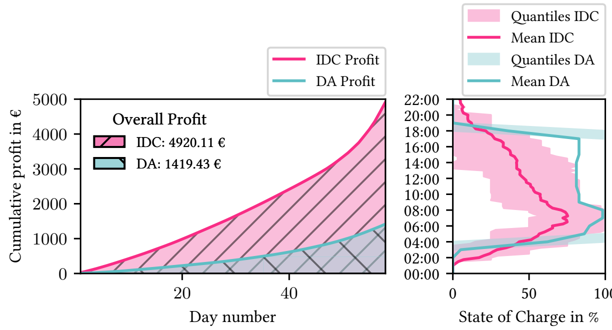
This section examines the proposed trading strategies, their corresponding BESS dispatch, and the associated profit potentials. To evaluate the general market potential, we choose to compare the overall profit made by the suggested bidding strategy. Comparing bidding strategies based on their profit potential is common in the literature [5, 52], as it determines the business case for BESS installations. We further enrich the analysis by analysing risk metrics and differences between single days. Through this analysis, we contribute to the broader discussion on multi-market BESS bidding strategies and how they compare to single-market approaches. In this section, we introduce the general rationale of the two main result plots before analysing each bidding strategy (single-market, multi-market: myopic and coordinated) in detail in the following sections.

The profitability of the trading strategies is illustrated using cumulative daily profits for all analysed algorithms, alongside aggregated charging and discharging decisions. Figure 2 presents the cumulative profits ordered by the daily profits per strategy on the left. Depending on whether single- or multi-market bidding is applied, the lines may either be stacked or remain separate. The right-hand plot depicts the average charging and discharging decisions over the evaluation period, providing a summary of dispatch results for each market. To further capture the variability in the resulting SoC dispatch patterns, the quantiles of the SoC are also visualised.

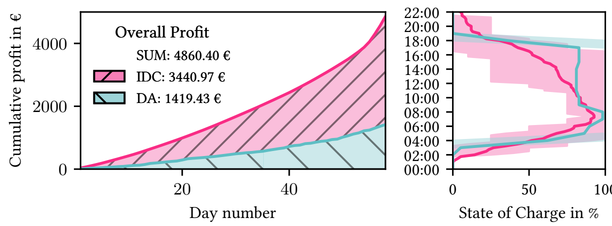
The empirical distribution of the profit is depicted in Figure 3. Taking the distribution of profits into account reveals risk differences. The figure quantifies the probability of reaching at least a certain profit during the model horizon. This plot also allows us to quantify the risk with the Conditional Value at Risk (CVaR). The CVaR is defined for a certain level α and quantifies the average profit that we make in the α-percentage of the worst cases and is used to quantify tail-risk. Since the analysed BESS trading is in general not causing losses, the CVaR is positive and, hence, a higher CVaR indicates lower tail-risk.

5.1 Single-Market Bidding

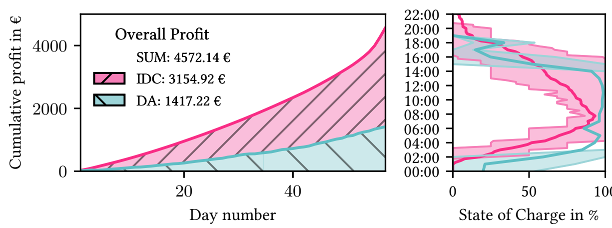
As a baseline, we analyse a single-market bidding strategy that independently participates in either DA or IDC trading. The results are presented in Figure 2a. For participation in the DA auction, a MILP approach is employed, utilising the EXAA price forecast. The blue line represents the cumulative profit over the model horizon. The relatively steady slope of this line indicates that DA profits remain stable across all modelled days. In contrast, the cumulative



(a) Single-Market Bidding: Individual MILP and RI



(b) Myopic Multi-Market Bidding: Stacked MILP and RI



(c) Coordinated Multi-Market Bidding: Stacked DRL and RI anticipating profits in RI

Figure 2: Overview of cumulative profit and SoC development over the model horizon, distinguished by the introduced three bidding strategies

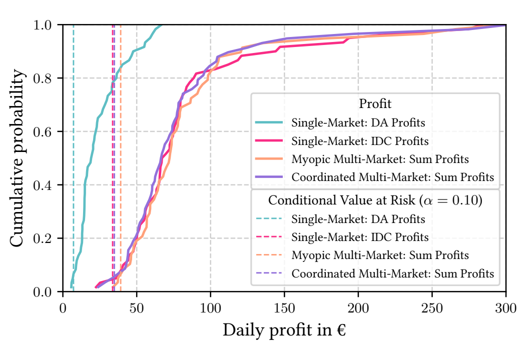


Figure 3: Empirical distribution function of the profits with horizontal lines indicating the CVaR

profit from participation in the IDC, shown by the pink line, exhibits a distinct increase in slope, particularly in the final 10 days. This suggests that profit opportunities in the IDC vary strongly across the model period. The shaded areas in the figure emphasise that these markets are modelled individually. The overall profit is shown at the end of the plotted model horizon (day 60) and is also written out in the plot area. As expected, the DA market generates considerably lower profits compared to the IDC, namely only 30% of the IDC profit. This result highlights the significant potential for BESS to exploit price fluctuations in the IDC, even when accounting for transaction costs and risk-free bidding.

Further insights into the resulting operational behaviour are provided by the SoC plots on the right side of the figure. The quantiles reveal that the DA charging and discharging behaviour is relatively consistent compared to the IDC charging and discharging behaviour. Charging typically occurs in the early morning, while discharging is concentrated in the evening, with only minor variations of about one hour between different modelled days. This behaviour aligns with relatively stable price forecasts used for the DA market. In contrast, participation in the IDC results in a more dynamic SoC profile. While the peak SoC is reached at similar times as in the DA market, the discharging behaviour exhibits greater variability. This indicates that price variations in the IDC differ significantly between days, leading to differing optimal discharging strategies. This observation is consistent with the variability in daily profits observed for the IDC.

Analysing the empirical distribution function further emphasizes the finding that particular days contribute most to the overall profit, as the IDC profit exhibits strong asymptotic behaviour at the end of the distribution. In comparison the DA profits are more stable. As previously shown in Figure 2, the overall profit generated in the DA market is lower than in the IDC market. In the IDC we can observe a distinct tail of the distribution, leading to a CVaR of 33 €/day.

5.2 Myopic Multi-Market Bidding

Evolving from single-market to multi-market bidding, we present the results for myopic bidding, as shown in Figure 2b. By definition (see Section 3.3), myopic bidding involves using the MILP optimisation for the DA market without considering the IDC during the DA decision-making process. Consequently, the results for the DA market are identical to those observed in the single-market case. However, in this case, the DA trades are passed to the RI, which subsequently makes bidding decisions for the IDC, leading to a combined overall profit.

In terms of profit development, the cumulative profit from the IDC exhibits a similar slope to the single-market case, with a notable steep increase during the final 10 days. However, the overall profit generated by the IDC is reduced compared to the single-market IDC bidding, namely by approximately 1500 €. Despite this reduction, the combined profit from both markets (DA + IDC) is nearly equivalent to the profit achieved in the single-market IDC case. It is only decreased by 60 €, which equals a reduction of 1.2%. This demonstrates that myopic multi-market participation does not increase overall profits beyond single-market participation in the IDC, again assuming no open positions between markets. However, it mitigates risks by shifting a portion of the trading volume to the earlier DA auction. This early participation allows profits to be secured without being fully exposed to the higher price volatility and unforeseen forecast errors that lead to price peaks in the IDC. This finding is further supported by the analysis of the CVaR in Figure 3. As shown by the dotted line, the myopic multi-market bidding increases the CVaR to approximately 40 €/day and, hence, reduces tail risk in comparison to the participation only in the IDC (CVaR 33 €/day).

The influence of the DA trades on IDC bidding behaviour can be observed in the deviations of the resulting SoC profile. Compared to the single-market case, the SoC resulting from IDC trading in the myopic scenario tends to remain at higher levels for longer periods during the day. Specifically, the discharging process does not begin immediately after the BESS is fully charged but is often delayed until the second price peak of the day, typically occurring in the evening. This behaviour highlights the impact of earlier DA trades on the flexibility and timing of IDC bidding decisions. The rolling nature of the RI algorithm can result in being constrained by prior trades made in the DA market, which may subsequently be subject to price increases over time. In contrast, the RI algorithm without earlier DA trades might select a different set of trades in the beginning. This alternative strategy could lead to a distinct and more favourable trajectory.

5.3 Coordinated Multi-Market Bidding

Now we switch from the MILP to the DRL approach and incorporate profit feedback from the IDC in the second multi-market bidding strategy, resulting in coordinated bidding (Figure 2c). Focusing first on the DA trading, we observe an overall profit comparable to the results from the previous analyses. This indicates that the DRL approach effectively handles DA trading, as seen in other literature. The cumulative profit shows a similar slope over the model horizon, suggesting stable daily profits. Comparing the SoC profiles resulting from DA trading across the two multi-market bidding strategies reveals that the DRL approach generally selects different charging and discharging hours. An analysis of the evolution of DA profiles during the training process shows that after the first step of the curriculum learning, where we only train on the DA market, a similar SoC pattern emerges, comparable to that observed in single-market bidding. However, including the IDC market in the second step of curriculum learning leads to the depicted more dispersed and earlier SoC profiles across different days. Since much of the IDC trading happens during the night, the DA charging trades are shifted to even earlier hours, which theoretically restricts the IDC trading less. Similarly, discharging happens earlier as well. This demonstrates that the DRL approach learned to adapt the DA profit potential while adhering to the technical constraints of the system in the designed environment.

Shifting the focus to the IDC, we observe only limited differences between the myopic multi-market approach. Based on expectations and previous results using stochastic optimisation [38], one would anticipate that the DRL approach learns to diversify its bids to anticipate and leverage IDC price developments. While the bids are indeed more diverse, they result in a reduction of the IDC profit by approximately 300 €, compared to myopic multi-market bidding. The resulting SoC profile after the IDC market is more volatile. This is as expected, since the DA charging profiles have a higher variability as well. This indicates a shift in the strategy toward earlier discharging behaviour, potentially driven by the altered profit dynamics in the IDC.

With regard to the analysis of the empirical distribution, we can observe that the coordinated multi-market bidding increases the tail risk in comparison to the myopic multi-market bidding, while still having slightly lower tail risk than the single-market participation in the IDC. The empirical distribution function suggests that the profit made by coordinated bidding is usually lower than made by the single-market IDC and myopic multi-market bidding, as indicated by the line being above and to the left of those two. However, there are cases in which coordinated multi-market bidding outperforms myopic bidding, which we describe in Appendix A.5. We observe that most days the difference between the myopic and coordinated bidding strategy is low, namely below 10% of the daily profit. Especially the earliest charging on the edge of the depicted quantile results in underperformance compared to myopic bidding. Furthermore, we observe that the difference between coordinated and myopic bidding increases (coordinated profit increases and/or myopic profit decreases) when lagged spreads $\overline{\mathrm{s}}_{d-1}$ are higher and variability in these spreads $\sigma_{s_{d-1}}$ is greater. Higher load $\mathrm{fc}_t^{\mathrm{load}}$ and wind forecasts $\mathrm{fc}_t^{\mathrm{wind}}$ as well as their daily variability $\sigma(\mathrm{fc}_d^{\mathrm{wind}})$ decrease the difference (coordinated profit decreases and/or myopic profit increases). Since we observe these relations, we expect the coordinated strategy to benefit from more information, namely, a larger observation space. This, however, would require more computational power.

6 Discussion

The findings of our study provide actionable insights for market participants, aiding in informed decisions regarding market participation and trading strategies. Firstly, even in a single-market case, relying solely on the DA market and neglecting the IDC market significantly limits the profit potential of BESS. This finding supports the relevance of BESS trading on the IDC and underlines the ability of the RI algorithm to model this potential, which has been only benchmarked against other IDC modelling strategies in previous literature [52]. We demonstrate that the risk-free RI algorithm on the IDC market achieves significantly higher profits than the DA market, even when realistic transaction costs are considered. Therefore, investment decisions for future BESS projects should incorporate these findings and account for IDC profit opportunities. Secondly, while myopic participation in both the DA and IDC markets does not increase profit potential compared to exclusive participation in the IDC market, it allows one-third of the profit potential to be secured a day in advance. This bidding strategy offers a means to mitigate trading risks effectively, which is evidenced by the reduction of tail risk.

The proposed coordinated multi-market bidding strategy does not achieve the full potential of market coordination observed in other studies. Studies using stochastic optimisations report a profit increase by coordinating across multiple markets, including the IDC of up to 28% for different storage types [38], and approximately 18% for a conventional and renewable portfolio, while simplifying the continuous nature of the IDC [41]. This raises an important further research avenue: Why does the presented strategy fail to replicate such behaviour?

Several reasons may account for this discrepancy. Firstly, the cited studies base their profit increases on artificial price scenarios, which may deviate from achievable real-world results. Additionally, they include ancillary service markets, which contribute to profits but remain secondary to spot markets. A more significant difference is the allowance of open positions between DA and IDC trades, enabling strategies like purchasing energy in the DA for later discharge in the IDC. While profitable, such strategies entail significant risk. In contrast, the presented approach reflects an implementation that disallows open positions. As a result, the combined use of MILP- and RI-based myopic multi-market bidding balances profit potential with reduced risk and modelling simplicity. Another possible explanation is that the DRL approach has not yet fully learned the required diversification strategies. While the coupling of IDC profits is evident, increased incentives for exploration and more powerful computational resources are needed to achieve better convergence. The latter would enable an extension of the observation space, which should increase the likelihood of anticipating IDC patterns. Beyond the scope of the introduced framework, we note that an increased number of agents pursuing the same trading strategies could influence market outcomes and ultimately canibalise profit potentials. We see this as an interesting avenue for further research, based on game-theoretic and agent-based models.

The chosen model horizon is selected for its relevance, as it reflects a period unaffected by the COVID-19 pandemic or the elevated electricity prices that began in 2021. However, it would be valuable to examine how coordination evolves with more recent data, particularly in light of higher price peaks that offer greater opportunities for BESS to capitalise on temporal arbitrage. Additionally, the recent decision to transition the DA auction to 15-minute intervals presents a compelling opportunity to explore our approach further. The implementable nature of the proposed framework ensures it can be easily adapted to incorporate updated market data, providing timely and actionable insights into emerging and potentially shifting market dynamics. Therefore, the proposed multi-market BESS framework serves as a valuable tool for practitioners and policymakers, offering practical insights for trading on spot electricity markets. Our open-source framework facilitates the modelling of BESS trading under uncertainty and realistic market conditions. By doing so, it supports more informed investment decisions and contributes to the integral expansion of BESS capacities.

Our open-source framework serves as a solid foundation for **further research** in the field of BESS trading algorithms, enabling modular advances. In conclusion, we see the following key topics as the most promising directions for future studies:

Mastering multi-market coordinated bidding: While we expected the coordinated multi-market bidding strategy, where expectations of potential IDC market profits influence DA market decisions, to outperform the myopic strategy, it did not. Although the DRL approach showed reasonable trading decisions and an overall learning process, it underperformed compared to the myopic approach, which also carries reduced risk compared to single-market bidding. Future improvements could focus on refining the DRL formulation, increasing the observation space or allowing risk by enabling open positions on the DA market that are closed on the IDC market. This would also require adapting the RI algorithm. Further, very recently, a speed-up version of the RI is introduced [48]. An incorporation of this approach should substantially enhance the learning process, leading to an expected better learning performance.

Integrating ancillary markets: Our study focuses exclusively on two key spot markets, but additional profit potential exists in reserve markets. The proposed framework allows for the inclusion of these ancillary markets in future studies, enabling an analysis of their impact on BESS profitability while still realistically implementing DA and IDC market trading.

7 Conclusion and Outlook

Multi-market bidding is essential for energy storage systems to maximise profitability by leveraging temporal price differences across the day-ahead and continuous intraday markets. This study evaluates and compares single-market and multi-market bidding strategies, including a myopic and coordinated approach, where the latter utilises deep reinforcement learning. Summing up the contributions of our study and setting it in relation to previous studies, we have contributed a comprehensive open-source framework to model and compare battery spot market trading strategies more realistically. In contrast, previous studies have rather focused on particular markets or pursued simplifications of continuous trading. The findings from our study, which models real-world battery trading, differ from results in more theoretical works [38, 40]. The results highlight that combining day-ahead and continuous intraday trading, even in a myopic framework, results in comparable profits to single-market participation while mitigating risk through early participation in the day-ahead market. However, the deep reinforcement learning-based coordinated bidding strategy falls short of fully capturing the potential profits observed in stochastic optimisation studies, emphasising the need for further improvements in bid diversification and learning efficiency.

In particular, in the scope of the realistic implementation, increased profits from multi-market bidding cannot be replicated, although reducing associated risks. This gives scope for further research. The recent shift to 15-minute intervals in the day-ahead auction and the increasingly volatile market environment provide opportunities to further test and refine the framework. Future work should focus on enhancing deep reinforcement learning strategies for better alignment with market conditions and exploring the integration of additional markets to fully realise the benefits of multi-market coordination.

Acknowledgments

Kim K. Miskiw thanks the German Federal Ministry for Economic Affairs and Climate Action (BMWK) for the funding of the ASSUME project under grant number BMWK 03EI1052A.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Grammarly and ChatGPT 4o for language refinement and grammar correction.

After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Harri Aaltonen, Seppo Sierla, Ville Kyrki, Mahdi Pourakbari-Kasmaei, and Valeriy Vyatkin. 2022. Bidding a Battery on Electricity Markets and Minimizing Battery Aging Costs: A Reinforcement Learning Approach. *Energies* 15, 14 (July 2022), 4960. https://doi.org/10.3390/en15144960

[2] René Aïd, Pierre Gruet, and Huyên Pham. 2016. An optimal trading problem in intraday electricity markets. *Mathematics and Financial Economics* 10 (2016), 49–85. https://doi.org/10.1007/s11579-015-0150-8

[3] Muhammad Anwar, Changlong Wang, Frits de Nijs, and Hao Wang. 2022. Proximal Policy Optimization Based Reinforcement Learning for Joint Bidding in Energy and Frequency Regulation Markets. In *2022 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 1–5. https://doi.org/10.1109/PESGM48719.2022.9917082

[4] Felix Baumgarte, Gunther Glenk, and Alexander Rieger. 2020. Business models and profitability of energy storage. *IScience* 23, 10 (2020).

[5] Ioannis Boukas, Damien Ernst, Thibaut Théate, Adrien Bolland, Alexandre Huynen, Martin Buchwald, Christelle Wynants, and Bertrand Cornélusse. 2021. A deep reinforcement learning framework for continuous intraday market bidding. *Machine Learning* 110, 9 (Sept. 2021), 2335–2387. https://doi.org/10.1007/s10994-021-06020-8

[6] Tom Brijs, Frederik Geth, Cedric De Jonghe, and Ronnie Belmans. 2019. Quantifying electricity storage arbitrage opportunities in short-term electricity markets in the CWE region. *Journal of Energy Storage* 25 (Oct. 2019), 100899. https://doi.org/10.1016/j.est.2019.100899

[7] Jun Cao, Dan Harrold, Zhong Fan, Thomas Morstyn, David Healey, and Kang Li. 2020. Deep Reinforcement Learning-Based Energy Storage Arbitrage With Accurate Lithium-Ion Battery Degradation Model. *IEEE Transactions on Smart Grid* 11, 5 (Sept. 2020), 4513–4521. https://doi.org/10.1109/TSG.2020.2986333

[8] Bruno Cárdenas, Lawrie Swinfen-Styles, James Rouse, Adam Hoskin, Weiqing Xu, and SD Garvey. 2021. Energy storage capacity vs. renewable penetration: A study for the UK. *Renewable energy* 171 (2021), 849–867.

[9] Aoxia Chen and Pankaj K. Sen. 2016. Deployment of battery energy storage system for energy arbitrage applications. In *2016 North American Power Symposium (NAPS)*. IEEE, Denver, CO, USA, 1–8. https://doi.org/10.1109/NAPS.2016.7747920

[10] Sridhar Chouhan, Deepak Tiwari, Hakan Inan, Sarika Khushalani-Solanki, and Ali Feliachi. 2016. DER optimization to determine optimum BESS charge/discharge schedule using Linear Programming. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, Boston, MA, USA, 1–5. https://doi.org/10.1109/PESGM.2016.7741576

[11] Andreas Coester, Marjan W Hofkes, and Elissaios Papyrakis. 2020. Economic analysis of batteries: Impact on security of electricity supply and renewable energy expansion in Germany. *Applied Energy* 275 (2020), 115364.

[12] Sumeyra Demir, Koen Kok, and Nikolaos G. Paterakis. 2023. Statistical arbitrage trading across electricity markets using advantage actor–critic methods. *Sustainable Energy, Grids and Networks* 34 (2023), 101023. https://doi.org/10.1016/j.segan.2023.101023

[13] Cao Di, Weihao Hu, Xiao Xu, Tomislav Dragičević, Qi Huang, Zhou Liu, Zhe Chen, and Frede Blaabjerg. 2020. Bidding strategy for trading wind energy and purchasing reserve of wind power producer – A DRL based approach. *International Journal of Electrical Power & Energy Systems* 117 (2020). https://doi.org/10.1016/j.ijepes.2019.105648

[14] Yi Dong, Zhen Dong, Tianqiao Zhao, and Zhengtao Ding. 2021. A Strategic Day-ahead bidding strategy and operation for battery energy storage system by reinforcement learning. *Electric Power Systems Research* 196 (July 2021), 107229. https://doi.org/10.1016/j.epsr.2021.107229

[15] EPEX SPOT. 2024. Trading Products | EPEX SPOT. https://www.epexspot.com/en/tradingproducts

[16] EPEX SPOT SE. 2024. EPEX SPOT – Fee Schedule.

[17] EPEX SPOT SE. 2024. Indices | EPEX SPOT. https://www.epexspot.com/en/indices

[18] EPEX SPOT SE. 2024. Trading Brochure | EPEX SPOT. https://www.epexspot.com/sites/default/files/download_center_files/21-03-09_Trading%20Brochure.pdf

[19] European Network of Transmission System Operators for Electricity. 2024. ENTSO-E Transparency Platform. https://transparency.entsoe.eu/

[20] EXAA Abwicklungsstelle für Energieprodukte AG. 2024. EXAA - EXAA Energy Exchange Austria with 5 auctions in AT, DE and NL. https://www.exaa.at/en/

[21] Benedikt Finnah, Jochen Gönsch, and Florian Ziel. 2022. Integrated day-ahead and intraday self-schedule bidding for energy storage systems using approximate dynamic programming. *European Journal of Operational Research* 301, 2 (Sept. 2022), 726–746. https://doi.org/10.1016/j.ejor.2021.11.010

[22] John Forrest, Ted Ralphs, Stefan Vigerske, Haroldo Gambini Santos, John Forrest, Lou Hafer, Bjarni Kristjansson, jpfasano, EdwinStraver, Jan-Willem, Miles Lubin, rlougee, a andre, jpgoncal1, Samuel Brito, h-i gassmann, Cristina, Matthew Saltzman, tosttost, and to st. 2024. *coin-or/Cbc: Release releases/2.10.12.* https://doi.org/10.5281/zenodo.13347261

[23] Josh Gray and Pankaj Khandelwal. 2004. Towards a realistic gas storage model. *Commodities Now* 7 (2004).

[24] Nick Harder, Anke Weidlich, and Philipp Staudt. 2023. Finding individual strategies for storage units in electricity market models using deep reinforcement learning. *Energy Informatics* 6, S1 (Oct. 2023), 41. https://doi.org/10.1186/s42162-023-00293-0

[25] F.-Javier Heredia, Marlyn D. Cuadrado, and Cristina Corchero. 2018. On optimal participation in the electricity markets of wind power plants with battery energy storage systems. *Computers & Operations Research* 96 (2018), 316–329. https://doi.org/10.1016/j.cor.2018.03.004

[26] Mehdi Jafari, Audun Botterud, and Apurba Sakti. 2022. Decarbonizing power systems: A critical review of the role of energy storage. *Renewable and Sustainable Energy Reviews* 158 (2022), 112077.

[27] Daniel R. Jiang and Warren B. Powell. 2015. Optimal Hour-Ahead Bidding in the Real-Time Electricity Market with Battery Storage using Approximate Dynamic Programming. *INFORMS Journal on Computing* 27, 3 (2015), 525–543. https://doi.org/10.1287/ijoc.2015.0640

[28] Katharina Johannsen. 2024. ÜNB verzeichnen 161 GW an Batterie-Anschlussanfragen. *Montel News* (October 29 2024). https://montelnews.com/de/news/8458570b-24b0-46c0-9fb4-e9c881de5660/uenb-verzeichnen-161-gw-an-batterie-anschlussanfragen Accessed: 2025-01-07, verified through personal communication with Transmission System Operators, 2024-11.

[29] Abraham Alem Kebede, Theodoros Kalogiannis, Joeri Van Mierlo, and Maitane Berecibar. 2022. A comprehensive review of stationary energy storage devices for large scale renewable energy sources grid integration. *Renewable and Sustainable Energy Reviews* 159 (2022), 112213.

[30] Timo Kern. 2019. Rückwirkungen von Batterie-Vermarktungsoptionen auf den Strommarkt. In *Proceedings of the 11 Internationale Energiewirtschaftstagung an der TU Wien (IEWT 2019), Vienna, Austria, 13–15 February 2019*.

[31] Gro Klæboe, Jørgen Braathen, Anders Eriksrud, and Stein-Erik Fleten. 2022. Day-ahead market bidding taking the balancing power market into account. *TOP* 30 (2022), 683–703. https://doi.org/10.1007/s11750-022-00645-1

[32] Aleksandra Komorowska, Piotr Olczak, Emil Hanc, and Jacek Kamiński. 2022. An analysis of the competitiveness of hydrogen storage and Li-ion batteries based on price arbitrage in the day-ahead market. *International Journal of Hydrogen Energy* 47, 66 (Aug. 2022), 28556–28572. https://doi.org/10.1016/j.ijhydene.2022.06.160

[33] Håkon Kongelf, Kristoffer Overrein, Gro Klæboe, and Stein-Erik Fleten. 2019. Portfolio size's effects on gains from coordinated bidding in electricity markets. *Energy Systems* 10, 3 (2019), 567–591. https://doi.org/10.1007/s12667-018-0294-8

[34] Emil Kraft, Marianna Russo, Dogan Keles, and Valentin Bertsch. 2023. Stochastic optimization of trading strategies in sequential electricity markets. *European Journal of Operational Research* 308, 1 (July 2023), 400–421. https://doi.org/10.1016/j.ejor.2022.10.040

[35] Kyung-Bin Kwon and Hao Zhu. 2022. Reinforcement Learning-Based Optimal Battery Control Under Cycle-Based Degradation Cost. *IEEE Transactions on Smart Grid* 13, 6 (Nov. 2022), 4909–4917. https://doi.org/10.1109/TSG.2022.3180674

[36] Malte Lehna, Björn Hoppmann, Christoph Scholz, and René Heinrich. 2022. A Reinforcement Learning approach for the continuous electricity market of Germany: Trading from the perspective of a wind park operator. *Energy and AI* 8 (2022), 100139. https://doi.org/10.1016/j.egyai.2022.100139

[37] Jinhao Li, Changlong Wang, Yanru Zhang, and Hao Wang. 2024. Temporal-Aware Deep Reinforcement Learning for Energy Storage Bidding in Energy and Contingency Reserve Markets. (2024). https://doi.org/10.48550/ARXIV.2402.19110 Publisher: arXiv Version Number: 1.

[38] Nils Löhndorf and David Wozabal. 2023. The Value of Coordination in Multimarket Bidding of Grid Energy Storage. *Operations Research* 71, 1 (Jan. 2023), 1–22. https://doi.org/10.1287/opre.2021.2247

[39] Thomas Mercier, Mathieu Olivier, and Emmanuel De Jaeger. 2023. The value of electricity storage arbitrage on day-ahead markets across Europe. *Energy Economics* 123 (July 2023), 106721. https://doi.org/10.1016/j.eneco.2023.106721

[40] Kim K Miskiw, Nick Harder, and Philipp Staudt. 2024. Multi Power-Market Bidding: Stochastic Programming and Reinforcement Learning. In *Proceedings of the 57th Hawaii International Conference on System Sciences | 2024*. 3062–3071. https://hdl.handle.net/10125/106753

[41] Kim K Miskiw, Emil Kraft, and Stein-Erik Fleten. 2025. Coordinated bidding in sequential electricity markets: Effects of price-making. *Energy Economics* 144 (2025), 108316. https://doi.org/10.1016/j.eneco.2025.108316

[42] Michał Narajewski and Florian Ziel. 2020. Econometric modelling and forecasting of intraday electricity prices. *Journal of Commodity Markets* 19 (Sept. 2020), 100107. https://doi.org/10.1016/j.jcomm.2019.100107

[43] Michał Narajewski and Florian Ziel. 2022. Optimal bidding in hourly and quarter-hourly electricity price auctions: Trading large volumes of power with market impact and transaction costs. *Energy Economics* 110 (2022), 105974. https:

//doi.org/10.1016/j.eneco.2022.105974
[44] Sanmit Narvekar, Bei Peng, Matteo Leonetti, Jivko Sinapov, Matthew E. Taylor, and Peter Stone. 2020. Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey. arXiv:2003.04960 [cs.LG] https://arxiv.org/abs/2003.04960
[45] Tomás Ochoa, Esteban Gil, Alejandro Angulo, and Carlos Valle. 2022. Multi-agent deep reinforcement learning for efficient multi-timescale bidding of a hybrid power plant in day-ahead and real-time markets. *Applied Energy* (2022). https://doi.org/10.1016/j.apenergy.2022.119067
[46] Stig Oedegaard Ottesen, Asgeir Tomasgard, and Stein-Erik Fleten. 2018. Multi market bidding strategies for demand side flexibility aggregators in electricity markets. *Energy* 149 (April 2018), 120–134. https://doi.org/10.1016/j.energy.2018.01.187
[47] Antonin RAFFIN, Quentin Gallouédec, Noah Dormann, Adam Gleave, Anssi, Alex Pasquali, Juan Rocamonde, M. Ernestus, Patrick Helm, Thomas Simonini, Quinn Sinclair, Corentin, Rohan Tangri, Sidney Tio, Tobias Rohrer, Tom Dörr, Wilson, Steven H. Wang, Sam Toyer, Roland Gavrilescu, Paul Maria Scheikl, Parth Kothari, Oleksii Kachaiev, Megan Klaiber, Marsel Khisamurdinov, Mark Towers, Jan-Hendrik Ewers, Grégoire Passault, Dominic Kerr, and Costa Huang. 2024. DLR-RM/stable-baselines3: Stable-Baselines3 v2.4.0: New algorithm (CrossQ in SB3-Contrib) and Gymnasium v1.0 support. https://doi.org/10.5281/ZENODO.8123988
[48] David Schaurecker, David Wozabal, Nils Löhndorf, and Thorsten Staake. 2025. Maximizing Battery Storage Profits via High-Frequency Intraday Trading. arXiv:2504.06932 [q-fin.TR] https://arxiv.org/abs/2504.06932
[49] Christian Schäfer. 2021. Alternative Erlösquellen für Batteriespeicher – Teil 1: Arbitrage im Stromhandel – Regelleistung Online. https://www.regelleistung-online.de/alternative-erlosquellen-fur-batteriespeicher-arbitrage-im-stromhandel/
[50] EPEX SPOT SE. 2024. All-time high volumes: Growth in spot markets illustrate trust of trading participants. https://www.epexspot.com/en/news/all-time-high-volumes-growth-spot-markets-illustrate-trust-trading-participants Accessed: October 14, 2024.
[51] Paul E. Seifert, Emil Kraft, Steffen Bakker, and Stein-Erik Fleten. 2024. Coordinated Trading Strategies for Battery Storage in Reserve and Spot Markets. arXiv:2406.08390 https://arxiv.org/abs/2406.08390
[52] Leo Semmelmann, Jannik Dresselhaus, Kim K Miskiw, Jan Ludwig, and Christof Weinhardt. 2024. An Algorithm for Modelling Rolling Intrinsic Battery Trading on the Continuous Intraday Market. In *Energy Informatics Review*, Vol. 4. 165–176. https://doi.org/10.1145/3717413.3717428
[53] Ana R. Silva, H.M.I. Pousinho, and Ana Estanqueiro. 2022. A multistage stochastic approach for the optimal bidding of variable renewable energy in the day-ahead, intraday and balancing markets. *Energy* 258 (2022), 124856. https://doi.org/10.1016/j.energy.2022.124856
[54] Qinghu Tang, Hongye Guo, and Qixin Chen. 2022. Multi-Market Bidding Behavior Analysis of Energy Storage System Based on Inverse Reinforcement Learning. *IEEE Transactions on Power Systems* 37, 6 (2022), 4819–4831. https://doi.org/10.1109/TPWRS.2022.3150518
[55] Vilayanur Viswanathan, Kendall Mongird, Ryan Franks, Xiaolin Li, Vincent Sprenkle, and R Baxter. 2022. 2022 grid energy storage technology cost and performance assessment. *Energy* 2022 (2022).
[56] Hao Wang and Baosen Zhang. 2018. Energy Storage Arbitrage in Real-Time Markets via Reinforcement Learning. In *2018 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, Portland, OR, 1–5. https://doi.org/10.1109/PESGM.2018.8586321
[57] David Wozabal and Gunther Rameseder. 2020. Optimal bidding of a virtual power plant on the Spanish day-ahead and intraday market for electricity. *European Journal of Operational Research* 280, 2 (2020), 639–655. https://doi.org/10.1016/j.ejor.2019.07.022
[58] Hanchen Xu, Xiao Li, Xiangyu Zhang, and Junbo Zhang. 2019. Arbitrage of Energy Storage in Electricity Markets with Deep Reinforcement Learning. https://doi.org/10.48550/ARXIV.1904.12232 Version Number: 2.
[59] Dimitrios Zafirakis, Konstantinos J. Chalvatzis, Giovanni Baiocchi, and Georgios Daskalakis. 2016. The value of arbitrage for energy storage: Evidence from European electricity markets. *Applied Energy* 184 (Dec. 2016), 971–986. https://doi.org/10.1016/j.apenergy.2016.05.047
[60] Bohan Zhang, Zhongkai Yi, Ying Xu, Jianing Xu, Yu Lu, and Yuhao Zhou. 2023. Expert Incorporated Deep Reinforcement Learning Approach for Market Arbitrage Strategy of the Battery Energy Storage. In *2023 IEEE 7th Conference on Energy Internet and Energy System Integration (EI2)*. IEEE, Hangzhou, China, 1709–1714. https://doi.org/10.1109/EI259745.2023.10513107
[61] Yuchen Zhou, Sarah Henni, and Philipp Staudt. 2022. Managing Intermittent Renewable Generation with Battery Storage using a Deep Reinforcement Learning Strategy. In *17. Internationale Tagung Wirtschaftsinformatik (WI 2022), Online*.
[62] Florian Ziel, Rick Steinert, and Sven Husmann. 2015. Forecasting day ahead electricity spot prices: The impact of the EXAA to other European electricity markets. *Energy Economics* 51 (Sept. 2015), 430–444. https://doi.org/10.1016/j.eneco.2015.08.005
[63] Tim Zonjee and Shahab Shariat Torbaghan. 2023. Energy Storage Arbitrage in Day-Ahead Electricity Market Using Deep Reinforcement Learning. In *2023 IEEE Belgrade PowerTech*. IEEE, Belgrade, Serbia, 1–7. https://doi.org/10.1109/PowerTech55446.2023.10202674

A Appendix

A.1 Literature Overview

Paper	Technology(s)	Market(s)	Timeframe: Case study	Algorithm	Episode length	Features in state space	Action space	Reward
Single-Market								
[63]	BESS	DA (FR)	Training: 2019 Testing: 2020	CDQN	24 steps	State of charge, price forecasts, hour counter	Discrete	Scaled revenue streams
[60]	BESS	DA (UK)	Training: 2019-2020 Testing: 2019-2020	DDPG	24 steps	State of charge, current price, price forecasts	Continuous	Scaled revenue streams
[56]	BESS	RT (US)	Training: 2016-2017 Testing: 2016-2017	Q-Learning	24 steps	Current price, state of charge	Discrete	Positive or negative reward based on average price
[35]	BESS	FCR (US)	n.a.	DQN	24 steps	State of charge, current price, frequency regulation signal	Discrete	Revenue streams, degradation costs
[7]	BESS	DA (UK)	Training: 2015 Testing: 2016	NN-DDQN, DQN, DDQN	24 steps	Price forecasts, state of charge	Discrete	Scaled revenue streams, degradation costs
[61]	BESS	DA (DE)	Training: 2018 Testing: 2019	DQN	24 steps	PV production forecasts, price forecasts, state of charge	Discrete	Revenue streams, service costs
[24]	BESS	DA (DE)	Training: 2019 Testing: 2019	MATD3	30 * 24 steps	Past & forecast residual load, past & forecast prices, past 6h state of charge, current energy charge costs	Continuous	Revenue streams
[1]	BESS	FCR (FI)	Training: 315 days of 2020 Testing: 35 days of 2020	REINFORCE, A2C, PPO	24 steps	Price forecast, hours since the battery rested last	Discrete	Revenue streams, penalties for market violations, battery aging
Multi-Market								
[37]	BESS	FCR & RT (AU)	Training: Jan-Okt 2016 Testing: Nov-Dec 2016	SAC	288 steps	State of charge, prices, extracted feature vector	Continuous	Sum revenues all markets
[14]	BESS	FCR & DA (US)	n.a.	FARL	24 steps	Clearing price of same hour last day, last decided bidding actions, State of charge, timestep	Discrete	Sum revenues both markets, cycling costs
[12]	None	BL & DA & IDC (NE)	Training: 2016-2019 Testing: 2020	A2C & A3C	Rolling monthly IDC contract counts	State of charge, prices and their quantiles, prices spreads, prices spreads, best quantities and prices	discrete (3)	Sum revenues IDC
[13]	Wind	BL & DA (DK)	Training: January 2016 - May 2018 Testing: May 2018 - October 2018	A3C	24 steps	wind forecast, last energy amount sold	continuous	Sum revenues IDC scaled to be $\in [-2, 0]$
[3]	BESS	BL & DA (AUS)	Training: 2016 Testing: first quarter 2017	PPO	24 steps	State of charge, spot market price, balancing market price	continuous	Sum revenues all markets
[45]	BESS & PV	RT & DA (AUS)	2017-2020	Two MVANNS	24 steps	Market prices and capacities, foregoing bids, solar infeed , ISO RT signals	continuous	Sum revenues all markets

Abbreviations: BESS = Battery Electric Storage System, PV = Photovoltaic, DA = Day-Ahead, IDC = Intraday Continuous, RT = Real-Time, FCR = Frequency Containment Reserve, BL = Balancing, ISO = Independent System Operator.

Table 1: Overview related literature focusing on single- and multi-market bidding using deep reinforcement learning

A.2 Example Results Rolling Intrinsic Algorithm

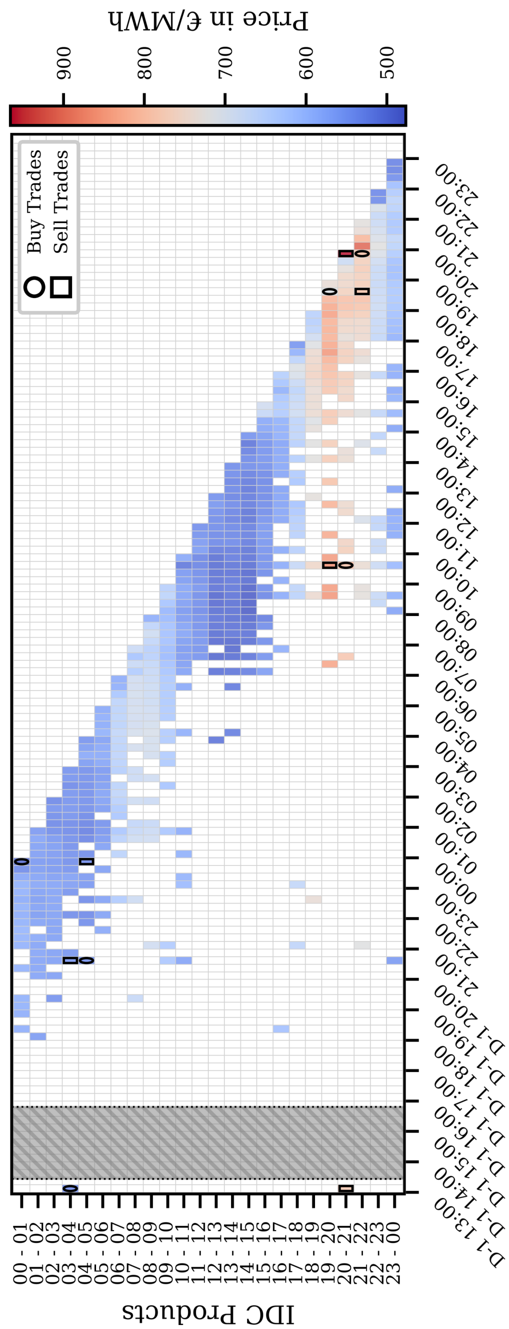


Figure 4: Trade executions per product of the RI algorithm for 2022-08-24 with DA positions

A.3 PPO Hyperparameters

Parameter	Value
λ	0.5
`policy`	ActorCriticPolicy
`learning_rate`	3e-3 to 3e-4 (linear decay)
`n_steps`	512
`batch_size`	64
`n_epochs`	10
`gamma`	0.99
`gae_lambda`	0.95
`clip_range`	0.4
`clip_range_vf`	None
`normalise_advantage`	True
`ent_coef`	0.05
`vf_coef`	0.4
`max_grad_norm`	0.5
`use_sde`	False
`sde_sample_freq`	-1
`rollout_buffer_class`	None
`rollout_buffer_kwargs`	None
`target_kl`	None
`stats_window_size`	100
`net_arch`	pi=[64, 64, 64, 64], vf=[64, 64, 64, 64]
`log_std_init`	-0.5
`verbose`	0
`seed`	42
`device`	`"CPU"`
`_init_setup_model`	True

Table 2: Parametrisation of the PPO model in Stable-Baselines3.

A.4 Technical Parameters

	Value	Source
C_{bess}	1.0 MWh	
P_{max}	1.0 MW	
η	0.86	[55]
Daily cycles	1.0	[32]
Transaction costs IDC	0.1 €/MWh	[16]

Table 3: Definition of the technical specification of the BESS in this case study.

A.5 Comparison Daily Profit

To analyse the differences between the myopic and coordinated multi-market bidding strategy, we plot the daily profit as a scatter plot in Figure 5. Each modelled day equals one plotted point, where the y-axis shows the coordinated multi-market profit and the x-axis is the myopic profit. Hence, every point above the 45°-line marks a day where we have a higher profit with coordinated multi-market bidding, and the point is hence colored green. A red point shows a day where the profit is higher with the myopic multi-market bidding strategy, respectively. Furthermore, we grey out all points that vary less than 10% from the 45°-line. Still, one can observe that many points are very close to the 45°-line, so in fact we do not observe high variance between the two approaches. An exception to this point is that they have a high overall daily profit.

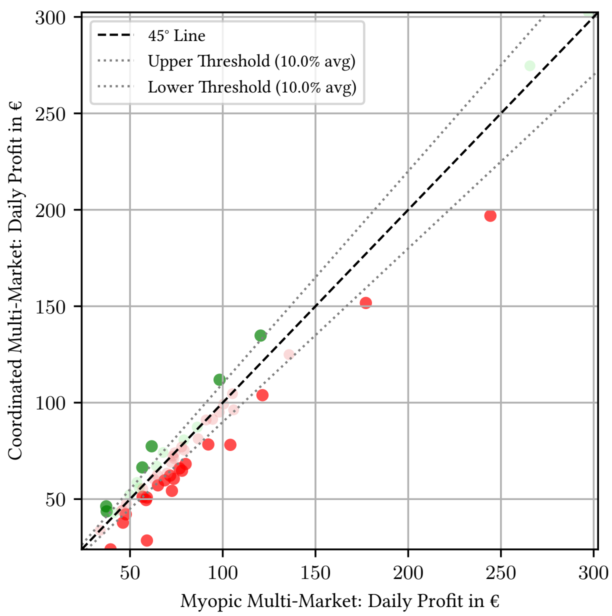


Figure 5: Comparison of myopic and coordinated daily profits highlighting absolute differences above 10% in comparison to the average daily profit of both multi-market bidding strategies.

Analysing the SoC profile for days exceeding the absolute variation of 10% reveals a general shift performed by the coordinated multi-market bidding strategy to earlier morning hours (beginning at 01:00 a.m.). Situations in which the coordinated bidding underperforms in comparison to myopic bidding are characterised by even earlier charging (finished by 01:00 a.m.) and sometimes discharging already before 06:00 a.m. As shown in Section 5.3, these correspond to the edges of the plotted quantiles. While both differentiate from the single-market DA bidding strategy, it seems that the algorithm lacks information to anticipate when further shifting to earlier morning hours is successful and when it is not.

To shed light on this, we carry out a correlation analysis between all potential features and the profit difference between coordinated and myopic multi-market bidding, which is shown in Figure 6. A negative correlation indicates that a higher value of this feature decreases the difference, leading to a comparably worse performance of coordinated bidding, and vice versa, a positive correlation indicates better performance of the coordinated approach. Generally, higher lagged spreads and a higher variation of the seen spreads lead to a better performance of coordinated bidding, while higher load forecasts and their variation, as well as a higher variation in the wind forecast, decrease the performance of coordinated bidding. As there is a correlation between the difference and the features present, we expect the coordinate approach to benefit from more information, namely, a larger observation space. This, however, would require more computational power.

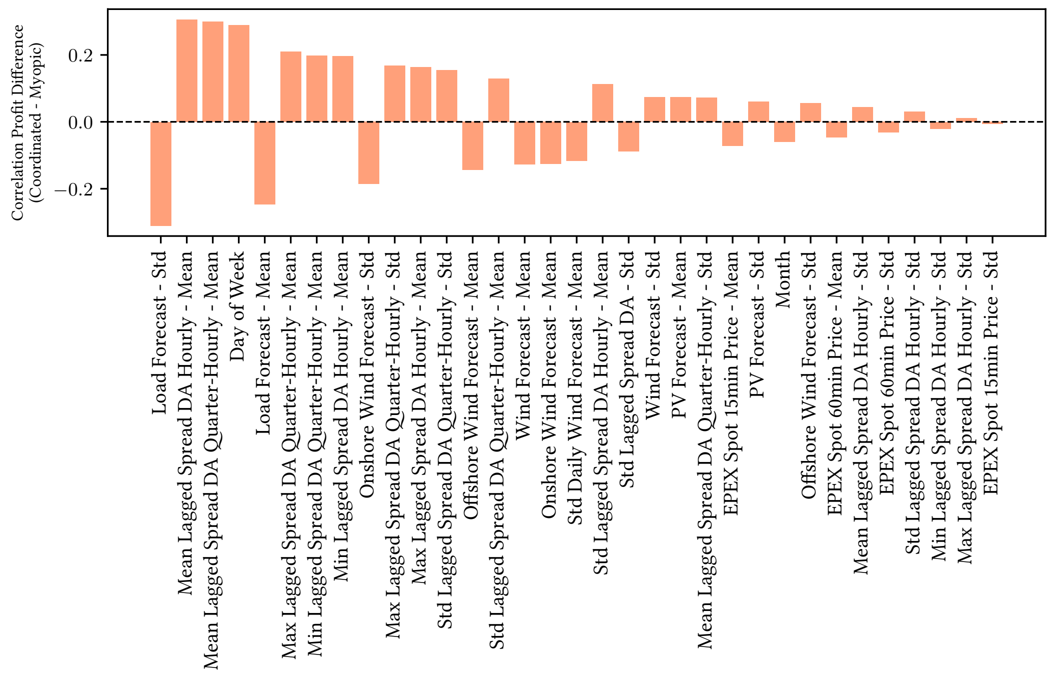


Figure 6: Correlation between the difference in daily profit of the coordinated and myopic multi-market bidding with hourly features available before DA clearing aggregated by their mean and standard deviation.