Modelling and Analysing the Impact of Local Flexibility on the Business Cases of Electricity Retailers

Imperial College London

Dawei Qiu

Supervisor: Prof. Goran Strbac

Department of Electrical and Electronic Engineering
Imperial College London

This dissertation is submitted for the degree of
Doctor of Philosophy

Control and Power Research Group March 2020

Declaration of Originality

As required by the college, I hereby declare that this thesis and the work described in it are my own work. Any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged through standard referencing practices of the discipline. No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Dawei Qiu
March 2020

Copyright Declaration

Dawei Qiu
March 2020

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Professor Goran Strbac, for providing me this opportunity to undertake this highly interesting research topic, his continuous support of my Ph.D. study and his expert guidance, wise advise, and immense knowledge over the last four years. His guidance helped me in all the time of research and writing of this thesis. Working with him has been an outstanding, fruitful and enjoyable experience in all aspects of academic life.

I would like to express my gratitude to my examiners, Professor Richard Vinter and Professor Anthony Papavasiliou, for a highly interesting discussion we had during the viva examination and for their valuable feedback on my thesis.

I am truly grateful to the Imperial College London for providing an outstanding research environment.

Besides my supervisor, I would like to thank my colleagues, Dr. Dimitrios Papadaskalopoulos and Dr. Yujian Ye, for giving me useful insights into this area, and helping me to grasp the underlying physical rationale behind the modelling and mathematics of this work.

I also appreciate the technical support and advice I gained from my colleagues Mingyang Sun, Temitayo Oderinwale, Jing Li from the Control and Power Research Group at the Imperial College London.

Thanks finally, from the bottom of my heart, to my deep loved father Guangxian Qiu and mother Lin Chen, who always believe in me, encourage me to follow my dreams and take great pride in my achievements. I am also grateful to my girlfriend Zichen Liang for her unconditional love, encouragement and companion. This Ph.D. study would not have been possible without them.

Abstract

Demand side response are proposed to incentivise customers to shift their electricity usage from peak demand periods to off-peak demand periods and to curtail their electricity usage during peak demand periods, which show great potential to reduce the peak loads, electricity prices, customers' bills and further stabilize the power systems. The investigation of this effect on the pricing strategies and the profits of electricity retailers has recently emerged as a highly interesting research area. However, the state-of-the-art, bi-level optimization modelling approach makes the unrealistic assumption that retailers treat wholesale market prices as exogenous, fixed parameters.

On the other hand, distributed energy resources (DER) in electricity markets are proposed to bring the significant operating flexibility which can support system balancing and reduce demand peaks, thereby limiting the balancing costs of conventional generators and the investments costs of new generation and network assets. And, local energy markets (LEM) have recently attracted great interest as they enable effective coordination of small-scale DER at the customer side, and avoidance of distribution network reinforcements. However, the introduction of LEM has also significant implications on the strategic interactions between the customers and incumbent electricity retailers, which has not been explored.

Furthermore, a specific demand response technology of electric vehicles (EV) exhibits the potential to support system balancing and limit demand peaks, thus improving significantly the cost-effectiveness of low-carbon electricity systems. And the effective pricing of EV charging by aggregators constitutes a key problem towards the realization of the significant EV flexibility potential in deregulated electricity systems and has been addressed by previous work through bi-level optimization formulations. However, the solution approach adopted in previous work cannot capture the discrete nature of the EV charging / discharging levels. Furthermore, aggregators suffering from communication and privacy limitations are hard to acquire the perfect knowledge of EV operating characteristics and traveling patterns.

Given such a context, this thesis aims at addressing the above challenges and proposing strategic retail pricing-based energy response programs to study the interactions between the electricity retailer / aggregator and its served flexible customers / EV based on game theoretic modeling and learning based approaches. We conduct the research in three different application scenarios:

1) This thesis proposes a novel bi-level optimization problem which represents endogenously the wholesale market clearing process as an additional lower-level problem, thus capturing the realistic implications of a retailer's pricing strategies and the resulting demand response on the wholesale market prices. This bi-level optimization problem is solved through converting it to a single-level Mathematical Programs with Equilibrium Constraints (MPEC). The scope of the examined case studies is threefold. First of all, they demonstrate the interactions between the retailer, the flexible consumers and the wholesale market and analyse the fundamental effects of the consumers' time-shifting flexibility on the retailer's revenue from the consumers, its cost in the wholesale market, and its overall profit. Furthermore, they analyse how these effects of demand flexibility depend on the retailer's relative size in the market and the strictness of the regulatory framework. Finally, they highlight the added value of the proposed bi-level model by comparing its outcomes against the state-of-the-art bi-level modelling approach.

2) This thesis explores for the first time the interaction between electricity retailer and LEM by proposing a novel bi-level optimization problem, which captures the pricing decisions of a strategic retailer in the upper-level problem and the response of both independent customers and the LEM (both including flexible consumers, micro-generators and energy storages) in the lower-level problems. Since the lower-level problem representing the LEM is non-convex, a new analytical approach is employed for solving the developed bi-level optimization problem. The examined case studies demonstrate that the introduction of an LEM reduces the customers' energy dependency on the retailer and limits the retailer's strategic potential of exploiting the customers through large differentials between buy and sell prices. As a result, the profit of the retailer is significantly reduced while the customers, primarily the LEM participants and to a lower extent non-participating customer, achieve significant economic benefits.

3) This thesis proposes a reinforcement learning (RL) method that the EV aggregator gradually learns how to improve its pricing strategies by utilizing experiences acquired from its repeated interactions with the EV and the wholesale market. Although RL can tackle the challenge of imperfect information and MPEC reformulation, the state-of-the-art RL methods require discretization of state and / or action spaces and thus exhibit limitations in terms of solution optimality and computational requirements. This thesis

proposes a novel deep reinforcement learning (DRL) method to solve the examined EV pricing problem, combining deep deterministic policy gradient (DDPG) principles with a prioritized experience replay (PER) strategy, and setting up the problem in multi-dimensional continuous state and action spaces. Case studies demonstrate that the proposed method outperforms state-of-the-art RL methods in terms of both solution optimality and computational requirements, and comprehensively analyze the economic impacts of smart-charging and vehicle-to-grid (V2G) flexibility on both aggregators and EV owners.

Table of contents

List of figures

List of tables

Nomenclature

The main notation used throughout the thesis is stated below for quick reference. Other symbols are defined as required throughout the text.

List of Abbreviations

DDPG	Deep deterministic policy gradient
DER	Distributed energy resources
DG	Duality gap
DIAG	Diagonalization algorithm
DNN	Deep neural network
DQN	Deep Q-network
DRL	Deep reinforcement learning
DSM	Demand side management
EPEC	Equilibrium problem with equilibrium constraints
ESS	Energy storage systems
EV	Electric vehicles
FD	Flexible demand
G2V	Grid-to-vehicle
GenCo	Generation company
KKT	Karush-Kuhn-Tucker

LEM	Local energy market
LL	Lower-level
MARL	Multi-agent reinforcement learning
MDP	Markov decision process
MG	Micro-generator
MILP	Mixed-integer linear programming
MIP	Mixed-integer programming
MIQP	Mixed-integer quadratic programming
MPEC	Mathematical programs with equilibrium constraints
PDDPG	Prioritized deep deterministic policy gradient
PER	Prioritized experience replay
RL	Reinforcement learning
TD	Temporal difference
TOU	Time-of-use
UL	Upper-level
V2G	Vehicle-to-grid
VOLL	Value of lost load

Nomenclature for Chapter 2

A. Indices and Sets

$t \in T$	Time periods
$b \in B$	Generation blocks of generation company
$\mathcal{R}$	Regulatory constraint set imposed on retail price
$c \in C$	Demand blocks of flexible consumer

$T^{gr} \in T$	Subset of periods when flexible EV is connected to the grid

B. Parameters

l^G	Linear cost coefficient of generation company / micro-generator (£/MWh)
q^G	Quadratic cost coefficient of generation company / micro-generator (£/MWh^2)
g^{max}	Maximum generation limit of generation company / micro-generator (MW)
λ_b^G	Marginal cost of block b of generation company (£/MWh)
g_b^{max}	Maximum generation limit of block b of generation company (MW)
λ_t^w	Wholesale price at period t (£/MWh)
l_t^D	Linear benefit coefficient of flexible consumer at period t (£/MWh)
q_t^D	Quadratic benefit coefficient of flexible consumer at period t (£/MWh^2)
d_t^{max}	Maximum demand limit of flexible consumer at period t (MW)
$\lambda_{t,c}^D$	Marginal benefit of block c of flexible consumer at period t (£/MWh)
$d_{t,c}^{max}$	Maximum demand limit of block c of flexible consumer at period t (MW)
s^{max}	Power capacity of energy storage (MW)
E^{min}	Minimum energy limit of energy storage / flexible EV (MWh)
E^{max}	Maximum energy limit of energy storage / flexible EV (MWh)
E^0	Initial energy level of energy storage / flexible EV (MWh)
η^c	Charging efficiency of energy storage / flexible EV
η^d	Discharging efficiency of energy storage / flexible EV
P^{max}	Maximum charging / discharging rate of flexible EV (kW)
E_t^{tr}	Energy requirements for travelling purposes of flexible EV at period t (kWh)

C. Variables

g_t	Generation of generation company / micro-generator at period t (MW)
$g_{b,t}$	Generation of block b of generation company at period t (MW)
λ_t^b	Retail price for buying energy from the retailer at period t (£/MWh)
λ_t^s	Retail price for selling energy from the retailer at period t (£/MWh)
d_t^w	Demand bid submitted by retailer in wholesale market at period t (MW)
d_t	Demand of flexible consumer at period t (MW)
$d_{t,c}$	Demand of block c of flexible consumer at period t (MW)
d_t^{sh}	Change of demand of flexible consumer at period t due to loading shifting (MW)
α	Load shifting limit of flexible consumer
s_t^c	Charging power of energy storage at period t (MW)
s_t^d	Discharging power of energy storage at period t (MW)
E_t	Energy level of energy storage / flexible EV at the end of period t (MWh)
V_t^c	Binary variable indicating whether flexible EV charges ($V_t^c = 1$) or not ($V_t^c = 0$) at period t
V_t^d	Binary variable indicating whether flexible EV discharges ($V_t^d = 1$) or not ($V_t^d = 0$) at period t
C_t	Charging power of flexible EV at period t (kW)
D_t	Discharging power of flexible EV at period t (kW)

D. Functions

$Cost_t^{GenCo}$	Quadratic cost function of generation company at period t (£)

$MCost_t^{GenCo}$ Linear marginal cost function of generation company at period t (£/MWh)

$Cost'^{GenCo}_{t,b}$ Cost function of block b of generation company at period t (£)

$MCost'^{GenCo}_{t,b}$ Marginal cost function of block b of generation company at period t (£)

$Pro_t^{Retailer}$ Profit function of electricity retailer at period t (£)

Ben_t^{FD} Quadratic benefit function of flexible consumer at period t (£)

$MBen_t^{FD}$ Linear marginal benefit function of flexible consumer at period t (£/MWh)

$Ben'^{FD}_{t,c}$ Benefit function of block c of flexible consumer at period t (£)

$MBen'^{FD}_{t,c}$ Marginal benefit function of block c of flexible consumer at period t (£/MWh)

Uti_t^{FD} Utility function of flexible consumer at period t (£/MWh)

$Cost_t^{MG}$ Quadratic cost function of micro-generator at period t (£)

Pro_t^{MG} Profit function of micro-generator at period t (£)

Pro_t^{EES} Profit function of energy storage at period t (£)

$Cost_t^{EV}$ Cost function of flexible EV at period t (£)

Nomenclature for Chapter 3

A. Indices and Sets

$t \in T$ Time periods

$i \in I$ Wholesale producers

$b \in B$ Generation blocks of wholesale producers

$c \in C$ Demand blocks of served consumers

B. Parameters

N_T	Length of market horizon
λ^{max}	Maximum limit of retail price (£/MWh)
K	Parameter determining the relation between average retail and wholesale prices
$\lambda^D_{t,c}$	Marginal benefit of block c of served consumers at period t (£/MWh)
$d^{max}_{t,c}$	Maximum demand limit of block c of served consumers at period t (MW)
α	Load shifting limit of retail demand
D^{bid}_t	Demand bid by other retailers in the wholesale market at period t (MW)
$\lambda^G_{i,b}$	Marginal cost of block b of wholesale producer i (£/MWh)
$g^{max}_{i,b}$	Maximum generation limit of block b of wholesale producer i (MW)
R^U_i	Ramp-up limit of wholesale producer i (MW)
R^D_i	Ramp-down limit of wholesale producer i (MW)
$g_{i,0}$	Initial generation of wholesale producer i (MW)
β	Relative size of the retailer in wholesale market

C. Variables

$d_{t,c}$	Demand of block c of served consumers at period t (MW)
d^{sh}_t	Change of demand of served consumers at period t due to loading shifting (MW)
d^{bid}_t	Demand bid submitted by retailer in wholesale market at period t (MW)
$g_{i,b,t}$	Generation of block b of wholesale producer i at period t (MW)
λ^r_t	Retail price at period t (£/MWh)
λ^w_t	Wholesale price at period t (£/MWh)

$\mu_{t,c}^{d-}, \mu_{t,c}^{d+}$	Dual variables associated with constraints (3.6) (£/MWh)
ξ	Dual variables associated with constraints (3.7) (£/MWh)
μ_t^{sh-}, μ_t^{sh+}	Dual variables associated with constraints (3.8) (£/MWh)
$\mu_{i,b,t}^{g-}, \mu_{i,b,t}^{g+}$	Dual variables associated with constraints (3.11) (£/MWh)
$\mu_{i,t}^{r-}, \mu_{i,t}^{r+}$	Dual variables associated with constraints (3.12) and (3.13) (£/MWh)

Nomenclature for Chapter 4

A. Indices and Sets

$t \in T$	Time periods
$i \in I$	Flexible consumers not participating in the LEM
$i' \in I'$	Flexible consumers participating in the LEM
$j \in J$	Micro-generators not participating in the LEM
$j' \in J'$	Micro-generators participating in the LEM
$k \in K$	Energy storages not participating in the LEM
$k' \in K'$	Energy storages participating in the LEM

B. Parameters

λ_t^w	Wholesale price at period t (£/MWh)
λ^{max}	Maximum limit of retail price (£/MWh)
$l_{i,t}^D$	Linear benefit coefficient of flexible consumer i at period t (£/MWh)
$q_{i,t}^D$	Quadratic benefit coefficient of flexible consumer i at period t (£/MWh^2)
$d_{i,t}^{max}$	Maximum demand limit of flexible consumer i at period t (MW)
l_j^G	Linear cost coefficient of micro-generator j (£/MWh)
q_j^G	Quadratic cost coefficient of micro-generator j (£/MWh^2)
g_j^{max}	Maximum generation limit of micro-generator j (MW)

s_k^{max}	Power capacity of energy storage k (MW)
E_k^{min}	Minimum energy limit of energy storage k (MWh)
E_k^{max}	Maximum energy limit of energy storage k (MWh)
E_k^0	Initial energy level of energy storage k (MWh)
η_k^c	Charging efficiency of energy storage k
η_k^d	Discharging efficiency of energy storage k

C. Variables

λ_t^b	Retail price for buying energy from the retailer at period t (£/MWh)
λ_t^s	Retail price for selling energy from the retailer at period t (£/MWh)
$d_{i,t}$	Demand of flexible consumer i at period t (MW)
$g_{j,t}$	Generation of micro-generator j at period t (MW)
$s_{k,t}^c$	Charging power of energy storage k at period t (MW)
$s_{k,t}^d$	Discharging power of energy storage k at period t (MW)
$E_{k,t}$	Energy level of energy storage k at the end of period t (MWh)
u_t	Binary variable indicating whether the LEM buys energy from the retailer ($u_t = 1$) or sells energy to the retailer ($u_t = 0$) at period t
n_t	Net demand of LEM at period t ($n_t > 0$ if the LEM buys energy from the retailer, $n_t < 0$ if the LEM sells energy to the retailer) (MW)
w_t	Net demand of retailer in the wholesale market at period t ($w_t > 0$ if the retailer buys energy from the wholesale market, $w_t < 0$ if the retailer sells energy to the wholesale market) (MW)

Nomenclature for Chapter 5

A. Indices and Sets

$t \in T$	Time periods

$i \in I$	Inflexible EVs
$j \in J$	Flexible EVs
$T_j^{gr} \in T$	Subset of periods when flexible EV j is connected to the grid

B. Parameters

Δt	Time resolution (h)
λ_t^w	Wholesale price at period t (pence/kWh)
λ^{min}	Minimum limit of retail price (pence/kWh)
λ^{max}	Maximum limit of retail price (pence/kWh)
$P_{i,t}^{inf}$	Charging power of inflexible EV i at period t (kW)
P_j^{max}	Maximum charging / discharging rate of flexible EV j (kW)
$E_{j,t}^{tr}$	Energy requirements for travelling purposes of flexible EV j at period t (kWh)
E_j^{min}	Minimum energy limit of flexible EV j (kWh)
E_j^{max}	Maximum energy limit of flexible EV j (kWh)
η_j^c	Charging efficiency of flexible EV j
η_j^d	Discharging efficiency of flexible EV j

C. Variables

λ_t^r	Retail price at period t (pence/kWh)
P_t^w	Net demand of aggregator in the wholesale market at period t (P_t^w > 0 if the aggregator buys energy from the wholesale market, P_t^w < 0 if the aggregator sells energy to the wholesale market) (kW)
$V_{j,t}^c$	Binary variable indicating whether flexible EV j charges ($V_{j,t}^c = 1$) or not ($V_{j,t}^c = 0$) at period t
$V_{j,t}^d$	Binary variable indicating whether flexible EV j discharges ($V_{j,t}^d = 1$) or not ($V_{j,t}^d = 0$) at period t

$C_{j,t}$	Charging power of flexible EV j at period t (kW)
$D_{j,t}$	Discharging power of flexible EV j at period t (kW)
$E_{j,t}$	Energy level of flexible EV j at period t (kWh)

List of Publications

1. Papers related to this thesis

Dawei Qiu, Dimitrios Papadaskalopoulos, Yujian Ye, and Goran Strbac. Investigating the impact of demand flexibility on electricity retailers. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7, 2018.

Dawei Qiu, Dimitrios Papadaskalopoulos, Yujian Ye, and Goran Strbac. Investigating the effects of demand flexibility on electricity retailers' business through a tri-level optimization model. *IET Gener. Transm. Distrib.*, Jan. 2020.

Dawei Qiu, Yujian Ye, and Dimitrios Papadaskalopoulos. Exploring the effects of local energy markets on electricity retailers and customers. *Electr. Pow. Syst. Res.*, Accepted.

Dawei Qiu, Yujian Ye, Dimitrios Papadaskalopoulos, and Goran Strbac. A deep reinforcement learning method for pricing electric vehicles with discrete charging levels. *IEEE Trans. Ind. Appl.*, Accepted.

2. Other papers

Dimitrios Papadaskalopoulos, Yujian Ye, Temitayo Oderinwale, and Dawei Qiu. A bi-level optimization modeling framework for investigating the role of flexible demand in deregulated electricity systems. In *19th International Conference on Environment and Electrical Engineering (19th IEEE EEEIC)*, 2019.

Yujian Ye, Dawei Qiu, Jing Li, and Goran Strbac. Multi-period and multi-spatial equilibrium analysis in imperfect electricity markets: A novel multi-agent deep reinforcement learning approach. *IEEE Access*, 7:130515–130529, Sep. 2019.

Yujian Ye, Dawei Qiu, Dimitrios Papadaskalopoulos, and Goran Strbac. A deep q network approach for optimizing offering strategies in electricity markets. In *2019 International Conference on Smart Energy Systems and Technologies (SEST)*, pages 1–6, 2019.

Yujian Ye, Dawei Qiu, Mingyang Sun, Dimitrios Papadaskalopoulos, and Goran Strbac. Deep reinforcement learning for strategic bidding in electricity markets. *IEEE Trans. Smart Grid*, 11(2):1343–1355, Mar. 2020.

Yujian Ye, Dawei Qiu, Xiaodong Wu, Goran Strbac, and Jonathan Ward. Model-free real-time autonomous control for a residential multi-energy system using deep reinforcement learning. *IEEE Trans. Smart Grid*, Early Access.

Chapter 1

Introduction

1.1 Context

Environmental and energy security concerns have driven governments worldwide to take significant initiatives towards the decarbonization of both generation and demand sides of energy systems [1]. However, these decarbonization initiatives introduce significant challenges to the operation and development of electricity systems. On the generation side, the decarbonization agenda involves the large-scale integration of renewable generation, which is however inherently characterized by high variability and limited controllability, challenging the cost-efficient balancing of the electricity system. On the demand side, the decarbonization agenda involves the electrification of transport and heat sectors, which is however expected to significantly increase demand peaks and drive capital-intensive generation and network investments.

In this setting, flexible demand technologies, enabling redistribution of electricity demand in time, have the potential to significantly improve the cost-effectiveness of low-carbon power systems by limiting demand peaks and increasing the use of renewable and cheaper generation sources. For example, at the demand side, the decarbonization agenda involves the electrification of certain sectors, with the electrification of the transport sector through the large-scale integration of electric vehicles (EV) being one of the key priorities. Amongst such flexible demand technologies, EV exhibit an outstanding flexibility potential due to their inherent ability to store electrical energy in their batteries, their low energy consumption requirements with respect to the significant capacity of their batteries, and the Vehicle-to-Grid (V2G) capability which enables EV to inject stored energy back to the grid. Numerous studies have investigated these beneficial impacts of different kinds of flexible demand technologies on electricity systems' short-term operation and long-term development [2–6].

One the other hand, a large number of small-scale distributed energy resources (DERs), including flexible loads, micro-generators and energy storages, are increasingly being connected to the distribution network, with the overall objective of providing the required flexibility to support the cost-effective development of low-carbon electricity systems. However, this paradigm change greatly complicates the operation of the system, as the effective coordination of such large numbers of DER involves very significant communication and computational scalability challenges as well as privacy concerns. In view of these challenges, local energy market (LEM) has recently emerged as an interesting approach to deal with these coordination challenges, as the global coordination burden is broken down to the coordination of local market clusters, each grouping a number of customers with DER, coordinating the energy exchanges between them and the upstream grid and addressing local network problems.

Beyond the above decarbonization initiatives however, governments worldwide have also taken significant initiatives towards the deregulation of the electricity industry, involving unbundling of vertically integrated monopoly utilities and the introduction of competition in both generation and retail sectors [7]. In this deregulated setting, beyond the above high-level impacts of demand side flexibility on the whole system (the whole society), it becomes imperative to investigate its impacts on the business case and strategies of different, self-interested market participants.

Another very important category of self-interested market participants in this deregulated setting includes electricity retailers (or suppliers, aggregators), especially when considering their direct interaction with the demand side. Specifically, these entities represent the large majority of the consumers in the wholesale electricity market, buying energy from this market at the wholesale prices and reselling it to their contracted consumers at certain retail prices. But these entities can also buy energy from its served customers with generation capacity at certain retail prices and resell to the wholesale market at the wholesale prices. The objective of these self-interested entities lies in maximizing their individual profits by optimizing the retail prices offered to their customers but also respecting regulatory limits [8]. By activating their flexibility, customers can respond to the offered retail prices so as to maximize their individual utility / profit. This effect will in turn impact the retailers' decisions regarding the offered prices and eventually their profit.

1.2 Motivation

A few recent works have modeled the dynamic interaction between the electricity retailer and the served customers through a bi-level optimization model. The upper-level problem represents the strategic decision making of the retailer who determines the optimal retail prices offered to the customers and the hourly demand / generation bids submitted to the wholesale market so as to maximize its profit over the considered daily horizon. The lower-level problem represents the decision making of customers which determine their optimal demand / generation response given the retail prices so as to maximize their total utility / profit.

However, all of them exhibit a fundamental shortcoming: the impact of the offered retail tariffs and the resulting energy response on the wholesale market is neglected. As a result, the wholesale market price is treated as an exogenous parameter. This assumption does not reflect the reality, as in most markets a relatively small number of retailers serve the whole population of consumers. In the UK for example, despite the early deregulation initiatives, the "Big Six" energy retailers still account for more than 75% of the market [9]. As a result, the modeling approaches employed in these works fail to fully capture the impact of customers' response on the retailer's strategies and profit.

As discussed above, the LEM concept addresses the communication and scalability challenges of DER coordination. However, the new LEM paradigm is also expected to have significant implications on the strategic interactions between the local end users and incumbent electricity retailers. With the introduction of LEM, the served customers together and exchange energy between each other, thus this paradigm change is expected to limit their dependency on the electricity retailers and reduce the volumes of energy traded by the retailers. This effect will in turn affects the retailer's decisions regarding the offered prices and eventually its profit.

Moreover, in the deregulated electricity sector environment, the realization of the EV flexibility potential needs to be integrated in electricity markets through an aggregator. Bi-level optimization model constitutes the most widely employed methodological framework in the existing literature for addressing this problem, where the upper-level problem represents the pricing optimization problem of the aggregator and the lower-level problems represent the response optimization problems of the EV. All relevant previous works have solved such bi-level optimization problems by converting them to single-level Mathematical Programs with Equilibrium Constraints (MPEC), through the replacement of the lower-level problems by their equivalent Karush-Kuhn-Tucker (KKT) optimality conditions.

Nevertheless, this solution approach exhibits two fundamental limitations. Firstly, it implicitly assumes that the aggregator has perfect knowledge of the EV operating characteristics and imperfect knowledge of the EV travelling patterns; such an assumption is not generally realistic, particularly when considering that the current penetration of EV is limited and thus existing work on the characterization of their operating characteristics (especially their travelling patterns) is far from comprehensive. Secondly, the lower-level problem does not include any binary decision variables since the derivation of the equivalent KKT optimality conditions is only possible when this problem is continuous and convex. As a result, previous works neglect physical non-convex operating characteristics of EV such as discrete charging / discharging levels, and therefore may lead to sub-optimal pricing strategies.

1.3 Research Questions

The research problem that we are looking at in this thesis focuses on strategic retail pricing problem of an electricity retailer who considers the effect of its served local flexibility. As an efficient strategic retail pricing strategy always considers both i) the interaction between the electricity retailer and the wholesale market, which affects its energy cost; and ii) the interaction between the electricity retailer and its served customers, which affects its energy revenue. In summary, there are three different scenarios when considering the potential questions in this thesis.

The first scenario happens when the electricity retailer serves a relatively large number of consumers, implying that its tariff strategy and the resulting demand response will have a significant impact on the wholesale market prices. The related research questions are: 1) how can we design a comprehensive framework to model the electricity retailer's strategic retail pricing, the served customers' demand response, and wholesale market clearing processing into an integrated fashion? 2) how can we design an efficient demand response program to model the consumers' characteristics of both time-independent demand elasticity and time-coupling demand shifting? 3) how can we model the regulatory constraints on the offered retail prices that prevent the retailer from exploiting the consumers and making excessive profits? 4) how can we analyse the impact of demand flexibility on electricity retailer's pricing strategies and the system parameters on its resulting business? 5) how can we evaluate the added value of our proposed model with respect to the state-of-the-art bi-level optimization model neglecting the interaction between the retailer's strategic pricing and the wholesale market conditions?

The second scenario happens when LEM is introduced in the retail market. The related questions are: 1) how can we design an efficient market mechanism to coordinate the energy exchanges between DER and the upstream grid? 2) how can we design an efficient framework to capture the interaction between retailer's strategic pricing and the LEM? 3) if the framework is designed as a bi-level optimization model, how can we solve this model in an effective way? This is because the lower-level problem of LEM is non-convex when it introduces the binary variables indicating whether the LEM buys energy from the retailer or sells energy to the retailer at the same time period. 4) how can we analyse the effects of introducing an LEM among the retailer's customers on the amount of energy served by the retailer, the buy and sell prices offered by the retailer, the retailer's profit and the different customers' economic surplus?

The third scenario happens when an aggregator represents a large number of EV in the wholesale market and coordinates their operation according to the market conditions and the EV operating characteristics. The related questions are: 1) how can we design a suitable mechanism for the EV aggregator to design effective time-specific prices, accounting for the discrete charging / discharging response patterns of the served EV? 2) how can we propose a new methodology to address the two fundamental limitations of MPEC approach as discussed in the previous section? 3) how can we address the curse of dimensionality of conventional Q-learning algorithms and accelerate the learning speed of deep reinforcement learning algorithms? 4) how can we demonstrate that the proposed method achieves a significantly higher profit for the examined aggregator and exhibits lower total computational requirements than the state-of-the-art reinforcement learning methods? 5) how can we analyse the impacts of smart-charging and V2G flexibility on the pricing decisions and profit of the aggregator as well as the costs of EV owners?

1.4 Contributions

This section outlines the contributions of this thesis, which are discussed from the problem development, model formulations, adopted methods, testing results and physical analysis to address the challenges associated with impacts of demand flexibility, local energy market on the electricity retailers as well as the strategic pricing problems for flexible electric vehicles charging. These contributions are further explored as follows:

1. *Investigating the Effects of Demand Flexibility on Electricity Retailers' Business through a Bi-Level Optimization Model*

- A bi-level optimization model is proposed, rigorously capturing the interactions between the retailer's pricing decisions (modeled in the upper-level), the flexible consumers' demand response (modeled in the first lower-level problem) and the wholesale market clearing process (modeled in the second lower-level problem). In contrast with state-of-the-art bi-level models, this model drops the unrealistic assumption that the retailer treats wholesale market prices as exogenous, fixed parameters, and represents endogenously the wholesale market clearing process; as a result, it captures the realistic implications of the retailer's pricing strategies on the wholesale market prices.

- Case studies with the proposed model demonstrate that demand flexibility reduces the retailer's revenue from the consumers, reduces the retailer's cost in the wholesale market, and does not have a uniform impact on the retailer's overall profit. This latter impact is shown to depend on the strictness of the regulatory constraints imposed on the offered retail prices, implying that demand flexibility can effectively complement regulatory policies in safeguarding the consumers against the strategic behavior of retailers.

- Case studies also demonstrate that the above effects of demand flexibility on the retailer's revenue and cost are relatively enhanced as the relative size of the retailer increases. However, the implications of the retailer's size on its overall profit are not uniform, but depend on the extent of demand flexibility, implying that new, small players in the retail market are more likely to take initiatives towards the realization of the flexibility potential of their consumers, than large, incumbent retailers.

- Finally, case studies demonstrate that the state-of- the-art bi-level models underestimate the above effects of demand flexibility on the retailer's revenue and cost, compared to the proposed model, and this underestimation is enhanced as the retailer's size increases. This result implies that state-of-the-art models are suitable for driving a retailer's decision-making only under the limiting condition that the retailer's size is extremely small (around 1% of the market according to the obtained results), in contrast with the general suitability of the proposed model.

2. *Exploring the Effects of Local Energy Markets on Electricity Retailers and Customers through a Bi-Level Optimization Model*

- A bi-level optimization model is proposed, rigorously capturing the interactions between the retailer's pricing decisions (modeled in the upper-level), the demand / generation response of flexible consumers, micro-generators, energy storages (modeled in the first three lower-level problems), and the operation of an local energy market among its served customers (modeled in the fourth lower-level problem). In contrast with state-of-the-art bi-level models on retailer's strategic pricing problems, this model makes the first attempt to fill the knowledge gap of modelling the interactions between the retailer's strategic pricing decisions and the operation of a local energy market among its served customers.

- Since this fourth lower-level problem is non-convex, as it includes the binary decision variables of the local energy market to either buy or sell energy to the retailer at each period, the formulated bi-level optimization problem cannot be solved through the traditional approach of converting it to a Mathematical Program with Equilibrium Constraints (MPEC). In this context, this thesis employs an approach, which is based on the relaxation and primal-dual reformulation of the non-convex lower-level problem and the penalization of the associated duality gap.

- Case studies with the proposed model demonstrate that the introduction of a local energy market is shown to reduce the customers' energy dependency on the retailer, since they are able to trade energy among them at prices which lie between the retailer's high buy prices and low sell prices, which is mutually beneficial for all flexible consumers, micro-generators and energy storages participants; regarding the latter, the local energy market is shown to unlock their arbitrage potential and activate them in the market.

- Case studies, consequently, also demonstrate that the retailer's strategic potential of exploiting the customers through large differentials between buy and sell prices is limited, and the retailer strives to make its offered buy and sell prices more competitive in order to attract more demand and generation by its customers. As a result of these effects, the profit of the retailer is very significantly reduced, while the customers enjoy significant economic benefits.

- Case studies demonstrate that this beneficial impact of local energy market is significantly higher for customers participating in the local energy market, but it is also substantial for non-participating customers, due to the above effects of the local energy market on the retailer's offered prices.

- Finally, case studies quantitatively demonstrate that the proposed relaxation method and primal-dual reformulation are effectively with the suitable selection of the penalization weighting factor in the final multi-objective optimization problem.

3. *A Deep Reinforcement Learning Method for Pricing Electric Vehicles with Discrete Charging Levels*

- A new bi-level optimization formulation is presented for modeling the examined EV pricing problem, which, in contrast with the existing literature, considers the V2G capability of EV and the discrete nature of their charging / discharging levels.

- A novel deep reinforcement learning (DRL) method is developed to solve the formulated problem, combining deep deterministic policy gradient (DDPG) principles with a prioritized experience replay (PER) strategy. In contrast with state-of-the-art reinforcement learning (RL) methods, this approach poses the examined problem in multi-dimensional continuous state and action spaces in a faster way.

- Case studies demonstrate that the proposed method achieves a significantly higher profit for the examined EV aggregator and exhibits lower total computational requirements than state-of-the-art RL methods.

- Case studies demonstrate that the proposed method exhibits a more favourable computational performance than benchmark RL methods due to the employment of the proposed PER strategy.

- Case studies apply the proposed method to different scenarios in order to comprehensively analyse the impacts of smart-charging and V2G flexibility on the pricing decisions and profit of the aggregator as well as the costs of EV owners.

1.5 Thesis Organization

The rest of this thesis is organized as follows:

Chapter 2 firstly introduces the different market mechanisms of wholesale market, retail market and local energy market in the deregulated electricity market and how the market participants interact with each other at different market mechanisms. Secondly, we provide the design and mathematical modeling of different market participants examined in this thesis. Finally, this chapter discusses the key assumptions made regrading the two proposed solution approaches of game-theoretic approach and learning-based approach, and presents the fundamental formulation for each of them adopted throughout the thesis.

Chapter 3 proposes a novel bi-level optimization model for the retail pricing decision making of a strategic electricity retailer, rigorously capturing the interactions between the retailer's pricing decisions (modeled in the upper-level), the flexible consumers' demand response (modeled in the first lower-level) and the wholesale market clearing process (modeled in the second lower-level). Such model is equivalent to an MPEC that can be recast as a tractable MILP problem using an exact linearization approach. Finally, the case studies with the proposed model demonstrate the impact of demand flexibility on the retailer's revenue from the consumers, its cost in the wholesale market, and its overall profit. This latter impact is shown to depend on the strictness of the regulatory constraints imposed on the offered retail prices and the relative size of the examined electricity retailer in the wholesale market. The add value of the proposed bi-level model is also evaluated by comparing its outcomes against the state-of-the-art bi-level modelling approach treating the wholesale market prices as exogenous, fixed parameters that are not affected by the consumers' response to the retail prices.

Chapter 4 explores for the first time the significant implications of local energy market on the strategic interactions between the customers and incumbent electricity retailers by proposing a novel bi-level optimization model, which captures the pricing decisions of a strategic retailer in the upper-level and the response of both independent customers and the local energy market (both including flexible consumers, micro-generators and energy storages) in the lower-level. Since the lower-level problem representing the local energy market is non-convex, the formulated bi-level optimization problem cannot be solved through the traditional approach of converting it to an MPEC adopted in Chapter 3. To this end, a new analytical approach based on the relaxation and primal-dual reformulation of the non-convex lower-level problem and the penalization of the associated duality gap is employed for solving the developed bi-level problem. The examined case studies comprehensively analyze the effects of introducing an local energy market among the retailer's customers on the amount of energy served by the retailer, the buy and sell prices offered by the retailer, the retailer's profit and the different customers' economic surplus.

Chapter 5 proposes a new bi-level optimization formulation for modeling the examined EV pricing problem, which, in contrast with the existing literature, considers the V2G capability of EV and the discrete nature of their charging / discharging levels. However, the game-theoretic solution approach adopted in Chapter 3 cannot capture the discrete nature of the EV charging / discharging levels and adopted in Chapter 4 assumes that the aggregator (solving the final MPEC) has perfect knowledge of the EV operating characteristics (which constitute parameters of the lower-level problems), such an assumption is not generally realistic, particularly when considering that the current penetration of EV is limited and thus existing work on the systematic characterization of their operating characteristics (especially their travelling patterns) is far from comprehensive. This chapter proposes a novel DRL method for designing effective retail prices by an EV aggregator, accounting for the discrete charging / discharging levels of the EV. This method is named prioritized deep deterministic policy gradient method (PDDPG), as it is founded on the combination of deep deterministic policy gradient (DDPG) principles and prioritized experience replay (PER) strategy. In contrast with previous RL methods, this method sets up the problem in multi-dimensional continuous state and action spaces. Case studies compare the performance of the proposed method against the two state-of-the-art methods (Q-learning and DQN) and investigate the economic impacts of EV flexibility on both the aggregator and the EV owners.

Chapter 6 concludes this thesis providing a summary, relevant conclusions drawn from the case studies carried out throughout the thesis work, and the main contributions of the thesis. Finally, some topics are suggested for the future research.

Appendix A provides the data of the electric vehicles used in Chapters 5.

Chapter 2

Market Models and Modelling Approaches

2.1 Market Mechanisms

In a traditional monopolistic or vertically integrated electricity market, market operators mainly aim to minimize the expected costs while maintaining an adequate security of supply [10]. Since 1980s, however, the electricity markets have been gradually evolving toward liberalized or deregulated structures, which are characterized by open competitive energy markets, unbundling electricity services, open access to the network, etc. To establish a competitive electricity market and improve its efficiency, the restructured market allows for exercising market power and tends to stimulate the emergence of new technologies [11, 12].

The key to open innovation in the deregulated power sector has been believed to be the development of consumer-centric business models and well-designed demand side management (DSM) programs [13, 14]. Following these ideas, the recent work in [15] looks even further forward to more subtle modeling of customer behavior, with considerations of their willingness to participate and even emotional or irrational features. With these prevailing ideas in the research community, the next-generation retail electricity market infrastructure will be a level playing field, where all energy end-users and customers have equal opportunity to play the role of active participants rather than pure passive price-takers [16, 8]. Fortunately, the recent development of the functionalities of the electricity retailers has opened many new possibility for monitoring, coordinating and controlling short-term delivery of electricity at the electricity retail side [17]. Especially with the further development of the concept of the demand side response, deregulation of the electricity market has been spreading out from wholesale market design into retail market design. In the new paradigm for energy transactions, different customers (e.g. flexible demand, micro-generation) or customer groups (e.g., energy community, local energy market) are strategical to affect the retailer's business as well as the wholesale market conditions.

Moreover, in smart grids, a growing number of customers will be able to have local generation capability, i.e., distributed energy resources (DERs), along with various flexible controllable loads, such as thermostatically-controlled loads (TCLs), distributed energy devices (DESs) and smart washing machines [18, 19]. Electric vehicles (EVs) and plug-in electric vehicles (PEVs) are also appealing as the most controllable loads because they can be curtailed for significant periods of time (e.g., several hours) without impact on end-use function [20, 21]. These kinds of customers are encouraged to actively participate in the retail market to provide local flexibility or localized power balance between energy surplus and energy deficit.

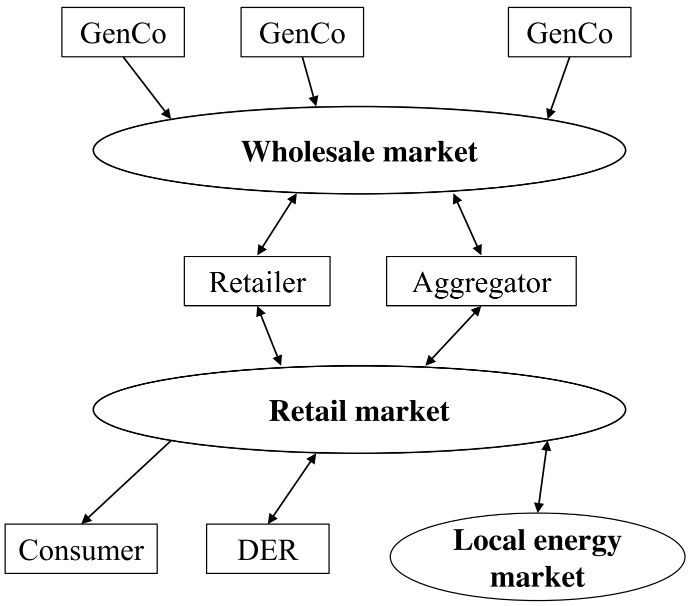

Fig. 2.1 The deregulation of the retail electricity market.

Fig. 2.1 shows a general perspective of energy interactions among different levels of power system decision makers in a deregulated retail electricity market. In this framework, generation companies (GenCos) are the first-level decision makers, electricity retailers (aggregators) are the second-level decision makers and end-customers (e.g., consumers, DERs, local energy market) are considered as the third-level decision makers. Other participants, due to their functionalities, may be located at each level of this framework. The first- and second-level decision makers are coupled with each other in the wholesale electricity market, which is managed by the market operator. The second- and third-level decision makers are coupled with each other in the retail electricity market, which is managed by the electricity retailers. Finally, part of end-customers (e.g. micro-generators and distributed energy storages) providing local generation capability are coupled with local demands that forms into a local energy market (LEM).

2.1.1 Wholesale Market

In this study, we model the wholesale electricity market as a pool-based, energy-only market and hourly resolution. This assumption captures essential features of the UK wholesale electricity market. The market employs a centralized market clearing mechanism, where the market operator receives the demand bids from the electricity retailers, derives the GenCos' cost function and their corresponding operational constraints, and determines the centralized solutions (constitute of market clearing dispatch of the GenCos and market clearing prices) by solving a generation cost minimization problem over the day-ahead horizon with hourly resolution.

Recently, there are a few papers [22–41] modeling the dynamic interaction between retailers and electricity consumers. In the retail market, by deploying demand flexibility, electricity consumers can modify their demand patterns according to the offered retail tariffs so as to maximize their utility. However, this effect on the wholesale market is neglected in these papers. As a result, the wholesale market price is treated as an exogenous parameter. This assumption does not reflect the reality, as in most markets a relatively small number of retailers serve the whole population of consumers. Therefore, each retailer serves a relatively large number of consumers, implying that its retail pricing strategy and the resulting demand response will have a significant impact on the wholesale market prices and consequently on the retailer's profit. As a result, the modeling approaches employed in these papers [22–41] fail to fully capture the impact of demand flexibility on the retailer's strategies and profit.

2.1.2 Retail Market

Retailers in the retail electricity market are supposed to purchase electricity in the wholesale electricity market and resell it to their subscribed end-user customers through assigning appropriate retail prices, either in a temporal variance way or at a flat rate.

Currently, the electricity retailer is usually operated as an entity that is independent of any generation or distribution company [8]. The decision-making process involved in buying and selling strategies usually contains some volatile market risks that are similar to the ones in any other market, such as the stock market and oil market. Especially with the further deregulation of the electricity market, along with the development of DSM and the proliferation of DERs, retailers participating in both the wholesale market and the retail market should carefully design their buying-selling trade-off and electricity portfolio optimization [42].

On the customer side, these end-consumers do respond to the retail pricing schemes offered by the retailers. Hence, the averaged flat retail prices under traditional designs give customers inaccurate information about the actual resources cost of power generation, and may ignore continuous changes in the electricity system conditions. Setting prices that differ for certain periods is an approach to realize these continuous market conditions. These time-based pricing schemes incentivise electricity customers to lower their usage during peak times, or shift their electricity usage from peak demand periods to off-peak demand periods. Time-of-use (TOU) tariffs play a crucial role in providing demand flexibility and tariff design in the electricity retail market [43]. Specifically, TOU tariffs are widely adopted by the retailer, which can achieve a better balance between realizing the demand flexibility and protecting consumers from pricing risk in the wholesale market. Among all the ToU tariffs, Economy 7 is adopted by UK electricity retailers to provide 7 hours of cheap off-peak electricity during the night. Prices during the rest of the time are, by contrast, relatively expensive [44].

However, TOU tariffs are fixed to consumers for several periods within the daily horizon (e.g. peak and off-peak hours) that cannot fully realize the flexibility potential of the consumers. Using TOU pricing schemes as baselines, some additional incentive mechanisms are also proposed on top of them to reflect the demand response from customers with energy awareness, which are aware of the electricity price elasticity and reasonable energy saving. In which, the time-specific retail pricing scheme is proposed in order to fully realize the flexibility potential of the consumers so as to develop the benefits of demand flexibility both on the market outcomes and the business of retailers. Overall, these kinds of negotiation-based demand response programs can be categorized as incentive mechanisms [45] that provide an additional economic management tool for the power system and market efficiency.

2.1.3 Local Energy Market

In most scenarios, customers play a role of energy consumption in retail electricity, purely serving as consumers of energy at the retail side. However, decentralisation constitutes one of the main features of the emerging Smart Grid. Specifically, a large number of small-scale DERs, including flexible loads, micro-generators and micro-storages, are increasingly being connected to the distribution network, with the overall objective of providing the much required flexibility to support the cost-effective development of low-carbon electricity systems [46]. However, this paradigm changes greatly complicates the operation of the system, as the effective coordination of such large numbers of DERs involves very significant communication and computational scalability challenges as well as privacy concerns, since DER owners are not generally willing to disclose private information and be directly controlled by external entities [47].

To develop strategies for these challenges, policy makers and planners need knowledge of how these DER can be integrated effectively and efficiently into a competitive retail electricity market. LEM has recently emerged as an interesting approach to deal with these coordination challenges, as the global coordination burden is broken down to the coordination of local market clusters, each grouping a number of customers with DER, coordinating the energy exchanges between them and the upstream grid and addressing local network problems. Beyond this coordination benefit, the local matching of power reduces net demand peaks and network losses, resulting in avoidance or deferral of capital-intensive network reinforcements.

However, beyond these system coordination and investment planning impacts, this new local market paradigm is expected to have significant implications on the business case of incumbent electricity retailers. The objective of these self-interested retailers lies in maximizing their individual profit by optimizing the offered electricity retail prices to the LEM, but also represent them in the wholesale electricity market by buying / selling energy consumed / produced by the LEM from / to the wholesale market. With the introduction of LEM, the DER group together and exchange energy between them. This paradigm change is expected to limit their dependency on the retailers and reduce the volumes of energy traded by the retailers. This effect will in turn impact the retailers' decisions regarding the offered retail prices and eventually their profits.

2.2 Market Participants

Having introduced the market mechanisms of wholesale market, retail market and local energy market in section 2.1, this section lies in providing the mathematical models of the market participants examined in this thesis.

2.2.1 Generation Companies

For presentation clarity reasons and without loss of generality, we assume that the Generation Company (GenCo) owns a single generation unit. And a simplified yet representative model is employed, where the quadratic cost function, linear marginal cost function and output limits of GenCo are expressed by (2.1), (2.2) and (2.3) respectively:

$$Cost_t^{GenCo}(g_t) = l^G g_t + q^G g_t^2 \tag{2.1}$$

$$MCost_t^{GenCo}(g_t) = l^G + 2q^G g_t \tag{2.2}$$

$$0 \leq g_t \leq g^{max}, \forall t \tag{2.3}$$

In order to avoid non-linearities in the optimization model of further Section 3.3, the quadratic cost function (2.1) is approximated by a piece-wise linear cost function, consisting of a number of generation blocks (segments) [48]. Since the cost curve of each segment is linear, the marginal cost of each segment is constant, which leads to a step-wise linear marginal curve. The cost, marginal cost function and output limits of each block b are expressed by (2.4), (2.5) and (2.6) respectively:

$$Cost'^{GenCo}_{t,b}(g_{t,b}) = \lambda_b^G g_{t,b} \tag{2.4}$$

$$MCost'^{GenCo}_{t,b}(g_{t,b}) = \lambda_b^G \tag{2.5}$$

$$0 \leq g_{t,b} \leq g_b^{max}, \forall t, \forall b \tag{2.6}$$

2.2.2 Electricity Retailers

A very important category of self-interested market participants in the deregulated electricity market includes electricity retailers, especially when considering their direct interaction with the demand side. Specifically, retailers represent the large majority of the consumers in the wholesale electricity market, buying energy from this market at the wholesale prices and reselling it to their contracted consumers at certain retail prices

[8]. The objective of a retailer lies in maximizing its individual profits by optimizing the retail buy prices λ_t^b and retail sell prices λ_t^s offered to its served customers and its demand bids / generation offers d_t^w in the wholesale market:

$$Pro_t^{Retailer}(\lambda_t^b, \lambda_t^s, d_t^w) = \lambda_t^b d_t - \lambda_t^s g_t - \lambda_t^w d_t^w \tag{2.7}$$

where d_t and g_t respectively expresses the demand and generation of the customers served by the retailer and λ_t^w indicates the wholesale prices.

It is assumed that customers are not allowed to directly buy / sell electricity from / to the wholesale market, as each customer is too small to participate in the wholesale market clearing. That means that the retailer is the only electricity provider who possesses strong market power. To alleviate the retailer's market power and build a fair retail market, the regulatory constraint set (2.8), such as the maximum and average retail prices (related to the wholesale prices), are subjected to prevent the retailer from exploiting the customers and making excessive profits [49]. Furthermore, the model can be enhanced by incorporating extra constraint set (2.8), (at the cost of extra complexity) to reflect better the true interaction between retailer and users or possible direct interaction between users and the wholesaler.

$$\lambda_t^b, \lambda_t^s \in \mathcal{R}, \forall t \tag{2.8}$$

Finally, the retailer has to ensure the energy sold to it served demand d_t, the energy bought from its served generation g_t and the energy traded in the wholesale market d_t^w are balanced at each time period t in (2.9).

$$d_t - g_t = d_t^w, \forall t \tag{2.9}$$

2.2.3 Distributed Energy Resources

In this study, a general framework for implementing a retail market is proposed as an electricity market structure with different distributed energy resources (DERs) penetration and demand side management of consumers. Moreover, the consumers are able to participate in the market as flexible demand, micro-generators and energy storage systems, but also a specific technology of electric vehicle.

Flexible Demand

Flexible Demand (FD) is based on the idea that the electricity use of consumers changes from their normal consumption patterns to the price of electricity over time. On the one hand, FD is used to induce lower electricity use at periods of high retail prices and higher electricity use at periods of low retail prices [50]. On the other hand, FD [51, 52] involves temporal redistribution of consumers' energy requirements. As a large number of researchers have stressed, consumers' flexibility regarding electricity use mainly involves shifting of their loads' operation in time instead of simply avoiding using their loads [13, 2]. In other words, load reduction during certain periods is accompanied by a load recovery effect during preceding or succeeding periods. This shift of energy demand from different periods drives a demand profile flattening effect.

Following the model employed in [53–55], the benefit obtained by the demand side at each time period is expressed through a quadratic, non-decreasing and concave function (2.10). The marginal benefit or "willingness to pay" is thus expressed through a linear decreasing function (2.11) which captures the effect of demand's self-price elasticity. As the demand level increases the consumers are willing to pay a lower price; equivalently, as the market price increases the demand requested by the consumers is reduced. The maximum price l_t^D that the consumers are willing to pay represents the value of lost load (VOLL) [56]. The limits in the requested demand level at each time period are expressed by (2.15). The VOLL, the slope of the marginal benefit function and the maximum demand limit are time-specific parameters, capturing the differentiated preferences of consumers across different time periods [57].

$$Ben_t^{FD}(d_t) = l_t^D d_t - q_t^D d^2 \tag{2.10}$$

$$MBen_t^{FD}(d_t) = l_t^D - 2q_t^D d \tag{2.11}$$

$$0 \leq d_t \leq d_t^{max}, \forall t \tag{2.12}$$

The quadratic benefit function (2.10) is approximated by a piece-wise linear benefit function, consisting of a number of blocks. The benefit, marginal benefit and demand limits of each block c are expressed by (2.13), (2.14) and (2.15) respectively:

$$Ben'^{FD}_t(d_{t,c}) = \lambda_{t,c}^D d_{t,c} \tag{2.13}$$

$$MBen'^{FD}_t(d_{t,c}) = \lambda_{t,c}^D \tag{2.14}$$

$$0 \leq d_{t,c} \leq d_{t,c}^{max}, \forall t, \forall c \tag{2.15}$$

The time-shifting flexibility of the demand side is expressed by (2.16)-(2.17). The variable d_t^{sh} represents the change of the demand with respect to the baseline level $\sum_c d_{t,c}$ at time period t due to load shifting, taking negative values when demand is moved away from t and positive values when demand is moved towards t. Constraint (2.16) ensures that demand shifting is energy neutral within the examined time horizon i.e. the total size of demand reductions is equal to the total size of demand increases (load recovery), assuming that demand shifting does not involve energy gains or losses. Constraint (2.17) expresses the limits of demand change at each time period due to load shifting as a ratio α $(0 \leq \alpha \leq 1)$ of the baseline demand; $\alpha = 0$ implies that the demand side does not exhibit any time-shifting flexibility, while $\alpha = 1$ implies that the whole demand can be shifted in time.

$$\sum_t d_t^{sh} = 0 \tag{2.16}$$

$$-\alpha \sum_c d_{t,c} \leq d_t^{sh} \leq \alpha \sum_c d_{t,c}, \forall t \tag{2.17}$$

The utility of the demand side at time period t is given by (2.18). While the energy payment (second term) depends on the final demand after any potential load shifting, the benefit (first term) is assumed to depend on the baseline demand; this assumption expresses the flexibility of the consumers to shift the operation of some of their loads without compromising the satisfaction they experience.

$$Uti_t^{FD}(d_{t,c}, d_t^{sh}) = \sum_c \lambda_{t,c}^D d_{t,c} - \lambda_t^b(\sum_c d_{t,c} + d_t^{sh}) \tag{2.18}$$

where λ_t^b expresses the retail buy prices offered by the retailer.

Distributed Generation

Micro-generators (MGs) are modern, small, on-site distributed energy generators that can operate grid-connected or be isolated from it [58]. They generally have capacities under 10 megawatts (MW) using renewable energy sources, such as solar panels and wind turbines, or high efficiency conversion of bio-energy or fossil fuels. Over the past decade there has been a strong push to accelerate the integration of micro-generators into the existing generation portfolio [59]. With suitable control they are able to reduce peak loads and can provide reliable power for commercial buildings, industrial facilities, and residential neighborhoods. What is more, they are attractive technologies in achieving specific local operational objectives, such as reliability, carbon emission reduction, diversification of energy sources, and cost reduction [60].

To this end, electricity retailers are key enablers in reaching these goals as they employ distributed generation and offer flexible energy management solutions. On the one hand, during the periods of supply-demand mismatch, micro-generators can interact among the demand side participants and trade electric generation over a marketplace. On the other hand, micro-generators make extra profit by selling their excess power to the electricity retailers. Hence, micro-generators derive benefits from both parties. In the literature, in order to quantify these benefits, the increasing quadratic cost functions are predominantly used to model the power generation cost of such generators. Then, the cost and generation limits are expressed by (2.19) and (2.20) respectively:

$$Cost_t^{MG}(g_t) = l^G g_t + q^G g_t^2 \tag{2.19}$$

$$0 \leq g_t \leq g^{max}, \forall t \tag{2.20}$$

where g_t is the power output at time period t, l^G and q^G are the cost coefficients of the generator, and g^{max} is the maximum generation limit of the generator. (2.19) usually serves as a part of the profit function (2.21) in the energy trading:

$$Pro_t^{MG}(g_t) = \lambda_t^s g_t - (l^G g_t + q^G g_t^2) \tag{2.21}$$

where λ_t^s is the retail sell price offered by the electricity retailer.

Energy Storage Systems

As the Energy Storage Systems (ESS) technology is becoming more economically viable, the role of ESS in energy trading will be more prominent. For large-scale renewable generation (e.g., solar arrays, wind farms), the ESS will be used to smooth out the output of the system [61]. On the other hand, for end-user applications, (distributed) community-based energy storage systems have already gained popularity [62]. In this case, the goal is to deploy small size storage units in the residential feeders to accommodate the demand of several houses during peak demand. Similarly for energy trading applications, the primary role of the energy storage system will be to store off-peak hour energy, so that users can use and exchange it during the periods of peak demand. Overall the goals of the ESS technologies are: 1) improving power grid optimization for bulk power production; 2) balancing the power system operations with intermittent renewable generation options; 3) providing ancillary services to grid operations.

The operational characteristics of an energy storage unit are expressed by (2.22)-(2.26). Constraint (2.22) expresses the energy balance in the storage unit including charging and discharging losses. Constraint (2.23) corresponds to its maximum depth of discharge and state of charge ratings. Constraints (2.24)-(2.25) represent its power limits. For the sake of simplicity, the storage energy content at the start and the end of the examined temporal horizon are assumed equal (2.26).

$$E_t = E_{t-1} + s_t^c \eta^c - s_t^d / \eta^d, \forall t \tag{2.22}$$

$$E^{min} \leq E_t \leq E^{max}, \forall t \tag{2.23}$$

$$0 \leq s_t^c \leq s^{max}, \forall t \tag{2.24}$$

$$0 \leq s_t^d \leq s^{max}, \forall t \tag{2.25}$$

$$E^0 = E_{NT} \tag{2.26}$$

The EES can be directly attached to the operating constraints of market mechanism, or is served by the electricity retailer with the objective of making arbitrage in price differentials:

$$Pro_t^{EES}(s_t^c, s_t^d) = \lambda_t^s \, s_t^d - \lambda_t^b \, s_t^c \tag{2.27}$$

where λ_t^s and λ_t^b are the retail sell and retail buy prices offered by the electricity retailer, respectively. It should be mentioned that the profit function (2.27) can be modified to take account of degradation cost of batteries, which is related to the depth of battery charge and discharge (2.23).

Electric Vehicles

Even though the primary goal of Electric Vehicles (EVs) is to offer environmentally friendly and cost-effective transportation options, the capability of EVs to store huge amount of electric power makes them a natural player in energy trading mechanism. With the use of bidirectional chargers, EVs can exchange electric power with the power grid or other market participants [63]. From energy trading standpoint, there are two emerging concepts on the use of EVs. The first one is Grid-to-Vehicle (G2V), in which the vehicle battery pack acts exactly the same as the ESS given in the previous section, but given the fixed energy requirement during the travelling. As a second scenario, which is also the most popular EV application concept is the Vehicle-to-Grid (V2G) where the stored energy is exchange with the grid. The predominant use of EVs in the literature is for energy trading to make extra profit by selling excess power and the related literature includes [64–66], however, the stored energy of EVs can also be used in ancillary services [67, 68, 68].

In literature, the existing works consider the coordination of EV charging based on the assumption that EVs can adjust their charging power continuously between zero and their maximum charging rates (i.e., continuous charging). However, due to the limitations of the current battery technology (e.g., the lithium-ion battery) and EV charger technology (e.g., the constant-current constant voltage approach [69]), EVs can only draw an approximately constant power during charging periods. Therefore, the binary variables V_t^c and V_t^d (2.28) indicate whether the EV charges ($V_t^c = 1$), discharges ($V_t^d = 1$), or remains idle ($V_t^c = V_t^d = 0$) at period t, with constraint (2.29) ensuring that charging and discharging cannot happen simultaneously. (2.30)-(2.31) express the fixed power rate (i.e., its maximum power rate at this period P_t^{max}) if the EV charges or discharges at period t.

$$V_t^c, V_t^d \in \{0, 1\}, \forall t \tag{2.28}$$

$$V_t^c + V_t^d \leq 1, \forall t \tag{2.29}$$

$$C_t = V_t^c \, P_t^{max}, \forall t \tag{2.30}$$

$$D_t = V_t^d \, P_t^{max}, \forall t \tag{2.31}$$

The parameter P_t^{max} is defined by (2.32); it is either equal to the power rating of the battery if the EV is connected to the grid at period t, or equal to zero if it is not (ensuring that charging or discharging cannot happen when the EV is not connected to the grid).

$$P_t^{max} = \begin{cases} P^{max}, & \text{if } t \in T^{gr}; \\ 0, & \text{otherwise.} \end{cases} \tag{2.32}$$

Constraints (2.33) express the energy balance in the EV battery, including the energy required for travelling purposes as well as charging and discharging losses. Constraints (2.34) express the minimum and maximum limits of the battery's energy content.

$$E_t = E_{t-1} + C_t \, \eta^c \, \Delta t - D_t / (\eta^d \, \Delta t) - E_t^{tr}, \forall t \tag{2.33}$$

$$E^{min} \leq E_t \leq E^{max}, \forall t. \tag{2.34}$$

Finally, the objective of the EV is minimizing the net cost of the EV charging / discharging, which is defined as the difference between i) its cost of buying energy from the retailer for charging (first term), and ii) its revenue from selling energy to the retailer through discharging (second term).

$$Cost_t^{EV}(C_t, D_t) = \lambda_t^b C_t - \lambda_t^s D_t \tag{2.35}$$

where λ_t^s and λ_t^b are the retail sell and retail buy prices offered by the electricity retailer, respectively.

2.3 Solution Approaches

The success of the energy trading mechanism heavily depends on the availability of the necessary communication infrastructures to ensure reliable information dissemination. In electricity market, participants need to update their demand or the amount of available energy to sell with the market place via two-way communication technologies. Also communication networks will enable trading entities to monitor their pricing information and available energy. The literature in its current state assumes that there is perfect communication between all players, then optimization based approach assuming perfect information and system knowledge is adopted to theoretically optimize the optimal decisions [70]. However, it is also important to quantify the impacts of communication system performance (e.g., privacy information, lack of system knowledge) on the operations of the energy trading mechanisms as it will create another level of uncertainty [71]. To this end, the data-driven learning based approach enables the trading entities to learn the optimal decisions by observing the limited system information due to the loss of knowledge.

The literature on electricity markets can be classified into several subcategories by considering the different combinations of employed enabling technologies that are presented in the previous section. Another important aspect in categorizing the literature is the employed modeling framework. In general, such frameworks can be classified into two categories. If the energy trading scenario is set to investigate one large-size player who tries to optimize its own utilities with considering the rest of the players and the market conditions, but also has the knowledge of the computational algorithm of the market mechanisms and the operating parameters of its interacting market players, in this case game theoretic approach is adopted to find the optimal solutions. However, due to the mathematical limitation and the computational issue, game theoretic approaches often demonstrate the difficulty of modeling the practical market in sufficient detail [72]. Simplified assumptions have to be made in the problem formulation in order to make the problem solvable. To this end, the models developed so far may fail to capture the operational characteristics of the system. Furthermore, in majority of the cases there are many uncertain behaviors and demand fluctuation in the market which is not more applicable in a realistic situation to acquire the knowledge of the system. In order to overcome the above drawbacks, the investigated market player adopts the learning-based approach to find the optimal solution by learning from the observed experiences.

2.3.1 Game-theoretic Approach

Bi-level optimization problem constitutes the most widely employed methodological framework of game-theoretic approach for developing the electricity market models over the last two decades [73]. The popularity of this methodology lies in its ability to capture the interaction between the strategic decision making of self-interested players (modeled in the upper-level as the leader) and the response of its interacting market players or the clearing of the electricity market (modeled in the lower-level as the follower).

If the lower-level problem is continuous and convex, this bi-level optimization problem can be solved by mathematical optimization software after converting it to a single-level *Mathematical Programs with Equilibrium Constraints* (MPEC), through the replacement of the lower-level problem by its equivalent optimality conditions [74]. The two main approaches for optimiality conditions are *Karush-Kuhn-Tucker* (KKT) optimality conditions and *Primal-Dual* transformation, where this conversion is illustrated in Fig. 2.2. In the first approach (KKT optimality conditions), a number of equalities are obtained from differentiating the corresponding Lagrangian with respect to the primal variables, and such equalities are equivalent to the set of primal and dual constraints of the second approach (primal-dual transformation). In addition, the set of complementarity conditions obtained by the first approach (KKT optimality conditions) is equivalent to the corresponding strong duality equality of second approach (primal-dual transformation) [75].

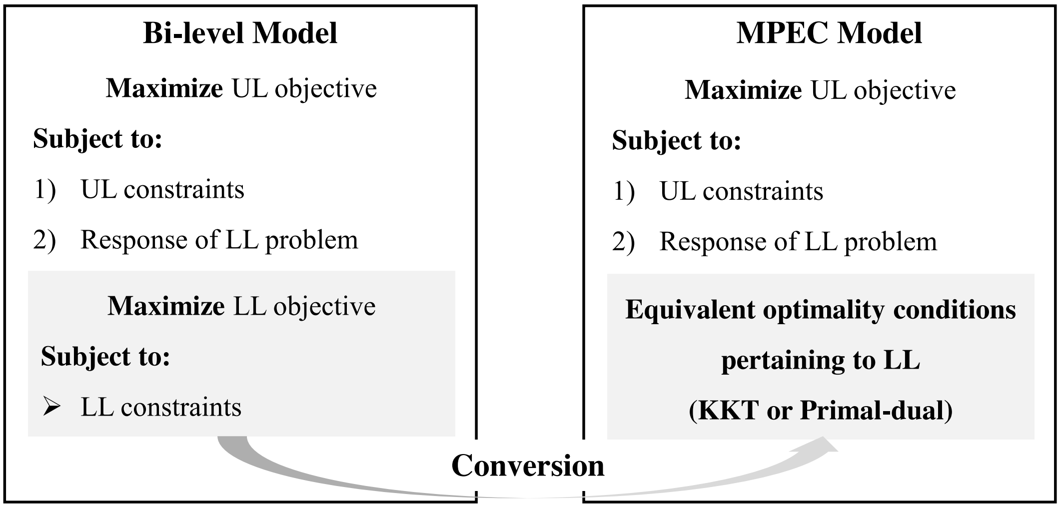

Fig. 2.2 Conversion of the bi-level model into its corresponding MPEC.

Formulation of the Bi-level Optimization Problem

This section presents the mathematical formulation for a general bi-level optimization problem. Moreover, different reactions could be generated in the lower-level towards each possible action conducted at the upper-level when multiple followers are involved in a bi-level decision-making. Hence, a general bi-level optimization problem can be easily extended into a bi-level multi-follower optimization problem, with one upper-level problem and multiple lower-level problems [76]. Mathematically, the objective function (2.36) of the upper-level problem is constrained by the upper-level equality and inequality constraints (2.37)-(2.38), and a set of N_l lower-level problems (2.39)-(2.41).

$$\min_{x} F(x, y_l^*) \tag{2.36}$$

subject to:

1. upper-level inequality and equality constraints:

$$G(x, y_l^*) \leq 0 \tag{2.37}$$

$$H(x, y_l^*) = 0 \tag{2.38}$$

2. a set of N_l lower-level problems:

$$y_l^* \in \operatorname*{argmin}_{y_l} \{ f_l(x, y_l) \tag{2.39}$$

subject to:

$$g_l(x, y_l) \leq 0 \tag{2.40}$$

$$h_l(x, y_l) = 0\}, \ \forall l. \tag{2.41}$$

The decision variables are divided into two sets, upper-level variables $x \in \mathbb{R}^{UL}$ and lower-level variables $y_l \in \mathbb{R}_l^{LL}$. The problem has two sets of constraints: the upper-level constraints (2.37)-(2.38) and the lower-level constraints (2.40)-(2.41). However, the problem has a hierarchical structure, as the upper-level problem includes the optimal solutions y_l^* to the lower-level problem, as indicated in the constraint (2.39). Instead, in the lower-level problem, the upper-level variables x are fixed parameters and not decision variables.

MPEC Reformulation with KKT optimality conditions

The KKT optimality conditions are a set of equality and inequality constraints that determine the optimal solutions of an optimization problem [77]. With necessary and sufficient KKT optimality conditions, the bi-level optimization problem (2.36)-(2.41) can be reformulated as an MPEC.

$$\min_{x, y_l, \mu_l, \lambda_l} F(x, y_l) \tag{2.42}$$

subject to:

1. upper-level inequality and equality constraints:

$$G(x, y_l) \leq 0 \tag{2.43}$$

$$H(x, y_l) = 0 \tag{2.44}$$

2. KKT optimality conditions associated with a set of N_l lower-level problems:

$$\{g_l(x, y_l) \leq 0 \tag{2.45}$$

$$h_l(x, y_l) = 0 \tag{2.46}$$

$$\mu_l \geq 0 \tag{2.47}$$

$$\nabla_{y_l} f_l(x, y_l) + \mu_l \nabla_{y_l} g_l(x, y_l) + \lambda_l \nabla_{y_l} h_l(x, y_l) = 0 \tag{2.48}$$

$$\mu_l g_l(x, y_l) = 0\}, \ \forall l. \tag{2.49}$$

where for each lower-level problem l, μ_l holds the dual variables to (2.40) and λ_l holds the free dual variables to (2.41). Note that we use the notation ∇_{y_l} for the gradient with respect only to the lower-level variables y_l. The constraints (2.49) are the complementary slackness constraints, making the MPEC a non-linearly constrained problem, irrespective of the original constraints.

The Non-linear complementarity constraints (2.49) are further handled by using the Fortuny-Amat mixed-integer reformulation [78] as presented below:

$$\min_{x,y_l,\mu_l,\lambda_l,\omega_l} F(x, y_l) \tag{2.50}$$

subject to:

$$(2.43) - (2.48) \tag{2.51}$$

$$\{\mu_l \leq \omega_l M_l^D \tag{2.52}$$

$$-g_l(x, y_l) \leq (1 - \omega_l)M_l^P \tag{2.53}$$

$$\omega_l \in \{0, 1\}\} \, \forall l. \tag{2.54}$$

where ω_l is additional binary variable and M_l^P, M_l^D are large enough constants. Model (2.50)-(2.54) is a mixed-integer linear optimization problem that can be solved using commercial software. And the objective function (2.50) is a bi-linear function of both variables x and y_l. However, the condition of problem (2.50)-(2.54) being equivalent to problem (2.42)-(2.49) when the large enough constants M_l^P, M_l^D are valid upper bounds for the primal and dual variables of the lower-level problem, respectively. Notice that appropriate values for M_l^P are often available, because they relate to primal variables, which are typically bounded by nature. However, M_l^D are upper bounds on dual variables and therefore, tuning these large enough constants is a more challenging task. The most commonly used trial-and-error tuning procedure [79].

MPEC Reformulation with Primal-Dual Optimality Conditions

To recast a bi-level optimization problem as an MPEC using primal–dual transformation approach, it is necessary to formulate the dual problems of the lower-level problems and the strong duality equalities. Consider that a set of N_l lower-level problems is formulated as below:

$$\min_{y_l} c^T x + d(x)^T y_l \tag{2.55}$$

subject to:

$$Ax + B(x)y_l \leq b_1 : \mu_l \tag{2.56}$$

$$Cx + D(x)y_l = b_2 : \lambda_l \tag{2.57}$$

$$y_l \geq 0 : \xi_l \tag{2.58}$$

where the dual variable μ_l, λ_l, ξ_l associated with the lower-level problems are indicated at the corresponding equations following a colon. Similarly to the bi-level problem (2.39)-(2.41), μ_l and λ_l are respectively the inequality and equality dual variables corresponding to the lower-level problem l. Additionally, dual variable ξ_l associates with the non-negativity of the primal variable y_l.

We can recast each lower-level problem l as an equivalent problem as:

$$c^T x + \min_{y_l} d(x)^T y_l \tag{2.59}$$

subject to:

$$B(x)y_l \leq b_1 - Ax : \mu_l \tag{2.60}$$

$$D(x)y_l = b_2 - Cx : \lambda_l \tag{2.61}$$

$$y_l \geq 0 : \xi_l \tag{2.62}$$

By using the duality theorems, the associated dual problem is:

$$c^T x + \max_{\mu_l, \lambda_l, \xi_l} \mu_l^T (b_1 - Ax) + \lambda_l^T (b_2 - Cx) \tag{2.63}$$

subject to:

$$B(x)^T \mu_l + D(x)^T \lambda_l = d(x) \tag{2.64}$$

$$\mu_l, \xi_l \geq 0 \tag{2.65}$$

$$\lambda_l : \text{free} \tag{2.66}$$

and the strong duality theorem is defined as:

$$d(x)^T y_l = \mu_l^T (b_1 - Ax) + \lambda_l^T (b_2 - Cx) \tag{2.67}$$

Considering the set of N_l lower-level primal problems (2.55)-(2.58), the set of their primal constraints (2.56)-(2.58), dual constraints (2.64)-(2.66) and the strong duality theorem (2.67) are equivalent to the KKT conditions and, therefore, they are sufficient conditions for the optimality.

Overall, the MPEC associated with the bi-level optimization problem (upper-level: (2.42)-(2.44), lower-level: (2.55)-(2.58)) is derived below by using the primal-dual optimality conditions.

$$\min_{x} F(x, y_l^*) \tag{2.68}$$

subject to:

1. upper-level inequality and equality constraints:

$$G(x, y_l^*) \leq 0 \tag{2.69}$$

$$H(x, y_l^*) = 0 \tag{2.70}$$

2. lower-level primal inequality and equality constraints:

$$Ax + B(x)y_l \leq b_1 \tag{2.71}$$

$$Cx + D(x)y_l = b_2 \tag{2.72}$$

$$y_l \geq 0 \tag{2.73}$$

3. lower-level dual equality and inequality constraints:

$$B(x)^T \mu_l + D(x)^T \lambda_l = d(x) \tag{2.74}$$

$$\mu_l, \xi_l \geq 0 \tag{2.75}$$

$$\lambda_l : \text{free} \tag{2.76}$$

4. lower-level strong duality theorem:

$$d(x)^T y_l = \mu_l^T (b_1 - Ax) + \lambda_l^T (b_2 - Cx) \tag{2.77}$$

2.3.2 Learning Based Approach

Reinforcement learning (RL) refers to a class of problems that are continuously learned from interaction with the environment and the methods to solve such problems. RL problems can be described as an agent (e.g., self-interested player in the upper-level problem) continuously learning from the interaction with the environment (e.g., the response of its interacting market players or the participated market in the lower-level problem) to achieve a specific goal, such as obtaining the maximum reward value (e.g., daily profit) [80]. However, one of the main challenges to RL is the need for manually designing quality features on which to learn. Each action cannot directly obtain the supervision information. It needs to be obtained through the final supervision information (reward) of the entire model, and it has a certain delay. In deep reinforcement learning (DRL) [81], deep neural networks are trained to approximate the optimal policy and/or the value function. In this way the DNN, serving as function approximator, allows the automatic discovery of features and enables powerful generalization.

RL Background

We now describe the background of reinforcement learning. In reinforcement learning, there are two objects that can interact: the agent and the environment.

1. **Agent** can sense the status of the external environment (State) and the reward of feedback (Reward), and learn and make decisions (Action). The decision-making function of the agent refers to making different actions according to the state of the external environment, and the learning function refers to adjusting the strategy according to the reward of the external environment.

2. **Environment** is everything outside the agent, and its state is changed by the action of the agent, and the corresponding reward is returned to the agent.

In RL, an agent acts within an environment by sequentially taking actions over a sequence of time steps $t \in T$, in order to maximize a cumulative reward. RL can be defined as a *Markov Decision Process* (MDP) which includes:

(a) a state space $\mathcal{S}$: a collection of the environment state;

(b) an action space $\mathcal{A}$: a collection of the agent's actions;

(c) a policy $\pi(a|s)$: a function of the agent to decide the next action according to the environmental state;

(d) a dynamics distribution with conditional transition probability $p(s_{t+1}|s_t, a_t)$, satisfying the *Markov property*, i.e. $p(s_{t+1}|s_t, a_t) = p(s_{t+1}|s_1, a_1, ..., s_t, a_t)$, represents the probability that the environment will change to the state $s^{'}$ at the next time step after the agent makes an action a according to the current state s;

(e) a reward r: $\mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$, that is, after the agent makes an action according to the current state s_t, the environment will give an immediate reward r_t to the agent, and this reward is related to the next state s_{t+1} after the action a_t.

The agent's decision in terms of which action a_t is chosen at a certain state s_t is driven by a policy $\pi(s_t) = a_t$. The agent deploys its policy to interact with the MDP and emit a trajectory of states, actions and rewards: $\tau = s_0, a_0, r_1, s_1, a_1, r_2, s_2, ..., s_{T-1}, a_{T-1}, r_T, s_T$ over $\mathcal{S} \times \mathcal{A} \times \mathbb{R}$. The agent starts from the perceived initial environment s_0, then decides to take a corresponding action a_1, the environment feeds back to the agent an instant reward r_1 and changes accordingly to the new state s_1, and then the agent makes one action a_1 according to state s_1, reward r_2 is rewarded and the environment is changed to s_2 accordingly. This interaction can continue until the end of the episode, illustrated in Fig. 2.3.

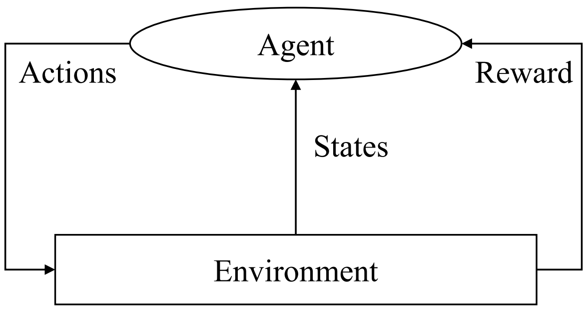

Fig. 2.3 Agent-environment interactions in RL.

RL Objective

Given the policy $\pi(a|s)$, the trajectory τ of an interaction process between the agent and the environment receives the cumulative reward (Return) $R(\tau) = \sum_{t=0}^{T-1} \gamma^t r_{t+1}$, which is the total discounted reward from time-step t onwards, where $\gamma \in [0, 1]$ is the discount factor which balances the importance between immediate and future rewards. When γ is close to 0, the agent is more concerned about short-term returns; when γ is close to 1, the long-term returns become more important.

Since both policy and state transition have a certain degree of randomness, the trajectory obtained by each experiment is a random sequence, and the total return the agent receives is also different. The goal of reinforcement learning is to learn a policy $\pi(a|s)$ that maximizes the cumulative discounted reward from the start state $t = 0$, i.e., the agent is expected to perform a series of actions to obtain as much average return as possible. The objective function of reinforcement learning is:

$$\mathcal{J}(\pi) = \mathbb{E}_{s\sim\rho^\pi}[R(\tau)] = \mathbb{E}_{s\sim\rho^\pi}[\sum_{t=0}^{T-1} \gamma^t r_{t+1}] \tag{2.78}$$

State-Action Value Function and Q-learning

To evaluate the expected return of policy π, we define *state-action value function* (or the *Q-value function*):

$$Q^\pi(s_t, a_t) = \mathbb{E}[R|s_t, a_t, \pi] \tag{2.79}$$

Q-value function constitutes an estimation of the expected, accumulative, discounted reward given an action a_t, at state s_t, and following the policy π from the succeeding states onwards. An optimal policy can be derived from the optimal Q-values $Q^*(s,a)$ $= \max_\pi Q^\pi(s,a)$ by selecting the action corresponding to the highest Q-value at each state.

To learn Q^π, we exploit the Markov property and define the function as a Bellman equation [82], which has the following recursive form:

$$Q^\pi(s_t, a_t) = \mathbb{E}_{s_{t+1}}[r_t + \gamma Q^\pi(s_{t+1}, \pi(a_{t+1}|s_{t+1}))] \tag{2.80}$$

The Bellman equation indicates that the action value function under the current policy can be decomposed in terms of itself. Namely, Q^π can be improved by bootstrapping, i.e., we can improve the estimate of Q^π by using the current estimate of Q^π through dynamic programming. This serves as the foundation of the Q-learning [83] algorithm, where the Q-values are updated as:

$$Q(s_t, a_t) \;\leftarrow\; Q(s_t, a_t) + \alpha\delta_t \tag{2.81}$$

$$\delta_t = r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \tag{2.82}$$

where $\alpha \in [0,1]$ is the learning rate and δ_t is the temporal difference (TD) error [84]. If the Q-value for each admissible state-action pair is visited infinitely often, and the learning rate α decreases over the time step t in a suitable way, then as $t \to \infty$, $Q^\pi(s,a)$ converges with probability one to $Q^*(s,a)$ for all admissible state-action pairs [83].

Deep Reinforcement Learning

Modeling policy $\pi(a|s)$ and state-action value functions $Q^\pi(s,a)$ are generally required in reinforcement learning. Early reinforcement learning algorithms focused on problems where states and actions were discrete and limited, and tables could be used to record these probabilities. But in many practical problems, the number of states and actions of some tasks is very large. For example, the chess game of Go has $3^{361} \approx 10170$ states, and the number of moves (i.e. positions of moves) is 361 [85]. There are also tasks whose states and actions are continuous. In order to effectively solve these problems, it is required to design a stronger policy function (such as deep neural network), so that the agent can deal with complex environments, learn better strategies, and have better generalization capabilities.

Deep Reinforcement Learning (DRL) is a combination of reinforcement learning and deep learning [86, 87]. Reinforcement learning is used to define problems and optimization goals, and deep learning is used to solve the modeling problems of policy and value functions. Deep reinforcement learning is a powerful, broadly applicable technique, that has be used to solve many complex problems not amenable to analytic solution or to other computational approaches. There are two different ways of combining reinforcement learning and deep learning, use deep neural networks (DNNs) to respectively model the state-action value function (value-based method) and the policy (policy-based method) in reinforcement learning, and then use the error back propagation algorithm to optimize the objective function.

In this thesis, we explore two popular deep reinforcement learning methods: 1) Deep Q-Network (DQN) method which employs a DNN to approximate the state-action value function, and has performed at the level of expert humans in playing Atari 2600 games [86] and demonstrated high quality performance in problems with continuous state spaces and discrete action space; 2) Deep Deterministic Policy Gradient (DDPG) method which concurrently learns a Q-function by the Bellman equation and uses the Q-function to learn a policy [88]. And it does so in a way which is specifically adapted for environments with both continuous state spaces and continuous action spaces.

Deep Q-Network

Deep Q-Network (DQN) leverages a Neural Network (parameterized by θ) to estimate the Q-value function, called Value Function Approximation.

$$Q_\theta(s,a) \approx Q^\pi(s,a) \tag{2.83}$$

If the actions are finitely discrete N_m actions $a_1, ..., a_m$, we can let the Q-Network input the environment state s and output an N_m-dimensional vector, where each dimension is represented by $Q_\theta(s, a_m)$, corresponding to the approximated value of value function $Q(s, a_m)$.

$$Q_\theta(s,a) = \begin{bmatrix} Q_\theta(s,a_1) \\ \vdots \\ Q_\theta(s,a_m) \\ \vdots \\ Q_\theta(s,a_{N_m}) \end{bmatrix} \approx \begin{bmatrix} Q^\pi(s,a_1) \\ \vdots \\ Q^\pi(s,a_m) \\ \vdots \\ Q^\pi(s,a_{N_m}) \end{bmatrix} \tag{2.84}$$

We need to learn a parameter θ so that the function $Q_\theta(s, a)$ can approximate the value function $Q^\pi(s, a)$. Taking Q-learning as an example, using temporal-difference (TD) learning method, let $Q_\theta(s, a)$ approximate $\mathbb{E}_{s_{t+1}}[r + \gamma\, Q_\theta(s_{t+1}, a_{t+1})]$, the objective (loss) function using gradient descent algorithm.

Prior to DQN, the employment of DNN for learning the Q-value function has generally been avoided since the learning process is prone to instability [87]. Nevertheless, the fact that DQN is able to learn the Q-value function using DNN in a stable and robust manner has been enabled by two innovations: the experience replay $\mathcal{R}$ and the target network $Q_{\theta'}(s, a)$ [87]. Concisely, the former method pools the collected experiences in a replay buffer and uniformly extracts samples to train the DNN, facilitating temporal de-correlation of consecutively generated training samples; the latter approach temporarily fixes the target Q-value during training, and thus stabilizes the learning process.

At each time step t, we sample a minibatch of N experiences $\{(s_n, a_n, r_n, s_{n+1})\}_{n=1}^{N}$ from $\mathcal{R}$, the training of the DNN is based on temporal difference (TD) learning [84] through the minimization of the following loss function, representing the mean-squared TD error:

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\Big(\underbrace{r_n + \gamma \max_{a_{n+1}} Q_{\theta'}(s_{n+1}, a_{n+1})}_{\text{target}} - \underbrace{Q_\theta(s_n, a_n)}_{\text{prediction}}\Big)^2 \tag{2.85}$$

and the following update is applied to the weights of DQN, where α is the learning rate for the gradient descent algorithm:

$$\theta \leftarrow \theta + \alpha \nabla_\theta \mathcal{L}(\theta). \tag{2.86}$$

Deep Deterministic Policy Gradient

Deep Deterministic Policy Gradient (DDPG) is a policy gradient algorithm that uses a stochastic behavior policy for good exploration but estimates a deterministic target policy, which is much easier to learn. Policy gradient algorithms utilize a form of policy iteration: they evaluate the policy, and then follow the policy gradient to maximize performance. DDPG is an actor-critic algorithm; it primarily uses two neural networks, one for the actor and one for the critic. These networks compute action predictions for the current state and generate a temporal-difference (TD) error signal each time step. The input of the actor network is the current state, and the output is a single real value representing an action chosen from a continuous action space. The critic's output is simply the estimated Q-value of the current state and of the action given by the actor. The deterministic policy gradient theorem provides the update rule for the weights of the actor network. The critic network is updated from the gradients obtained from the TD error signal.

The weights θ^Q of the critic network can be updated with the gradients obtained from the loss function in (2.85). The actor network is updated with the Deterministic Policy Gradient. The objective is to learn a deterministic policy $\mu_{\theta^\mu}(s)$ which gives the action that maximizes $Q_{\theta^Q}(s,a)$. But rather than globally maximizing $Q_{\theta^Q}(s,a)$, the critic provides gradients $\nabla_a Q_{\theta^Q}(s,a)$ which suggest directions of change of action leading to higher estimated Q-values. To calculate the policy loss, we further take the derivative of the objective function with respect to the policy parameter and apply the chain rule:

$$\nabla_{\theta^\mu} J(\mu_{\theta^\mu}) = \nabla_a Q_{\theta^Q}(s,a)\nabla_{\theta^\mu}\mu_{\theta^\mu}(s) \tag{2.87}$$

Similar as DQN, we update the policy with batches of experience and target network, we take the mean of the sum of gradients calculated from the mini-batch:

$$\nabla_{\theta^\mu} J(\mu_{\theta^\mu}) = \frac{1}{N}\sum_{n=1}^{N} \nabla_a Q_{\theta^Q}(s,a)|_{s=s_n,a=\mu_{\theta^\mu}(s_n)}\nabla_{\theta^\mu}\mu_{\theta^\mu}(s)|_{s=s_n} \tag{2.88}$$

and the following update is applied to the weights of the actor network, where α is the learning rate of the gradient decent algorithm:

$$\theta^\mu \leftarrow \theta^\mu + \alpha\nabla_{\theta^\mu} J(\mu_{\theta^\mu}) \tag{2.89}$$

Chapter 3

Strategic Pricing for Demand Flexibility

3.1 Introduction

Environmental and energy security concerns have driven governments worldwide to take significant initiatives towards the decarbonization of both generation and demand sides of energy systems [1]. However, these decarbonization initiatives introduce significant challenges to the operation and development of electricity systems. At the generation side, the decarbonization agenda involves the large-scale integration of renewable generation, which is however inherently characterized by high variability and limited controllability, challenging the cost-efficient balancing of the electricity system. At the demand side, the decarbonization agenda involves the electrification of transport and heat sectors, which is however expected to significantly increase demand peaks and drive capital-intensive generation and network investments.

In this setting, flexible demand technologies, enabling temporal redistribution (shifting) of electricity demand in time, have attracted great interest. This is because demand flexibility exhibits the potential to support system balancing and reduce peak demand levels, contributing to a more cost-effective transition to the low-carbon future. Numerous studies have investigated these beneficial impacts of demand flexibility on electricity systems' short-term operation and long-term development [2–6].

Beyond the above decarbonization initiatives however, governments worldwide have also taken significant initiatives towards the deregulation of the electricity industry, involving unbundling of vertically integrated monopoly utilities and the introduction of competition in both generation and retail sectors [7]. In this deregulated setting, beyond the above high-level impacts of demand flexibility on the whole system (the whole society), it becomes imperative to investigate its impacts on the business case and strategies of different, self-interested market participants. For example, authors in [13, 53, 89] have investigated the impacts of demand flexibility on the strategic behavior and profits of large electricity producers.

Another very important category of self-interested market participants in this deregulated setting includes electricity retailers (or suppliers), especially when considering their direct interaction with the demand side. Specifically, these entities represent the large majority of the consumers in the wholesale electricity market, buying energy from this market at the wholesale prices and reselling it to their contracted consumers at certain retail prices. The objective of these self-interested entities lies in maximizing their individual profits by optimizing the retail prices offered to their consumers but also respecting regulatory limits [8]. By activating their demand flexibility, consumers can respond to the offered retail prices so as to maximize their individual utility. This effect will in turn impact the retailers' decisions regarding the offered prices and eventually their profit.

3.2 Literature Review

Numerous previous works have investigated the interactions between flexible consumers and electricity retailers [22–41]. Table 3.1 summarizes in a structured way the main characteristics of these works. The representation of consumers' flexibility in these papers includes the elasticity to modify their overall energy requirements and / or the ability to shift their energy requirements in time. Furthermore, these papers generally adopt two different methodologies to model the interaction between an electricity retailer and its flexible consumers.

Specifically, papers [22–29] employ single-level optimization models, which aim at maximizing the retailer's profit; in these models, the optimal response of the flexible consumers to the retail prices is expressed as closed-form price-demand functions. The drawback of this approach lies in the fact that the parameters of these functions are determined based on exogenous data and therefore cannot accurately capture the impacts of the retail prices on the consumers' response.

In order to address this limitation, papers [30–41] employ bi-level optimization models in order to rigorously capture the interactions between the optimization of the

Table 3.1 Summary of existing literature associated with the examined problem.

Paper	Optimization model	Demand flexibility model	Wholesale market model	Conclusions around impact of demand flexibility on retailer's business
[22]	Single-level	Elasticity	Exogenous prices	Reduced revenue and profit
[23]	Single-level	Elasticity	Exogenous prices	Reduced revenue and profit
[24]	Single-level	Elasticity and shifting	Exogenous prices	Reduced revenue and cost, increased profit
[25]	Single-level	Elasticity and shifting	Exogenous prices	Reduced revenue and profit
[26]	Single-level	Elasticity	Exogenous prices	Reduced revenue, increased profit
[27]	Single-level	Elasticity and shifting	Exogenous prices	Increased revenue, cost and profit
[28]	Single-level	Elasticity and shifting	Exogenous prices	-
[29]	Single-level	Elasticity and shifting	Exogenous prices	-
[30]	Bi-level (MPEC)	Shifting	Exogenous prices	Reduced revenue, cost and profit
[31]	Bi-level (MPEC)	Shifting	Exogenous prices	Increased revenue, cost and profit
[32]	Bi-level (MPEC)	Elasticity and shifting	Exogenous prices	-
[33]	Bi-level (MPEC)	Elasticity	Exogenous prices	-
[34]	Bi-level (MPEC)	Elasticity and shifting	Exogenous prices	Elasticity reduces profit, shifting increases it
[35]	Bi-level (MPEC)	Elasticity and shifting	Exogenous prices	-
[36]	Bi-level (iterative)	Shifting	Exogenous prices	Reduced revenue, cost and profit
[37]	Bi-level (iterative)	Elasticity and shifting	Exogenous prices	Reduced revenue and cost
[38]	Bi-level (MPEC)	Elasticity and shifting	Exogenous prices	Increased profit
[39]	Bi-level (iterative)	Elasticity and shifting	Exogenous prices	-
[40]	Bi-level (MPEC)	Elasticity and shifting	Exogenous prices	Increased profit
[41]	Bi-level (MPEC)	Shifting	Exogenous prices	Reduced profit
This work	Bi-level (MPEC)	Elasticity and shifting	Endogenous clearing	(See Section 3.5)

pricing decisions of the self-interested retailer (modeled in the upper level – UL) and the optimization of the demand response of its flexible consumers (modeled in the lower level – LL), which is thus represented endogenously and not based on exogenous data. These bi-level optimization problems are solved either through converting them to Mathematical Programs with Equilibrium Constraints (MPEC) or by employing iterative approaches (Table 3.1).

Despite this methodological improvement achieved in [30–41], all previous relevant papers [22–41] exhibit another fundamental shortcoming. The prices in the wholesale market (from which the retailer buys energy) are treated by the retailer as exogenous, fixed parameters that are not affected by the consumers' response to the retail prices. This assumption is only valid when considering a retailer serving a very small population of consumers in the market. However, it does not generally reflect the reality, as in most countries a relatively small number of retailers serve the largest share of the market. In the UK for example, despite the early deregulation initiatives, the "Big Six" energy retailers still account for more than 75% of the market [9]. Therefore, considering a large retailer, its pricing strategies and the resulting demand response of its consumers will have a significant impact on the wholesale market prices. As a consequence, its pricing strategies will also affect its profit, given that the latter depends on both its revenue from the consumers as well as its cost in the wholesale market. In other words, the modeling approaches employed in previous works fail to comprehensively capture the effects of demand flexibility on the retailer's business case.

Beyond this methodological shortcoming, these previous works seem to provide conflicting conclusions regarding the overall impact of demand flexibility on the retailer's business (Table 3.1). While papers [22, 23, 25, 30, 36, 41] indicate that demand flexibility is likely to reduce the retailer's profit, papers [24, 26, 27, 31, 38, 40] indicate the exact opposite trend. We believe that these conflicting conclusions are driven by a combination of the methodological shortcoming of these works and their insufficient analysis of different parameters of the problem.

3.3 Approach

This chapter aims at addressing these limitations of previous works investigating the effects of demand flexibility on electricity retailers. In order to address their methodological shortcoming, this chapter proposes a novel bi-level optimization model capturing the interactions between the retailer's pricing decisions, the flexible consumers' demand response and the wholesale market clearing process in an integrated fashion. This bi-level optimization problem is efficiently solved after converting it to an MPEC, and subsequently to a Mixed-Integer Linear Problem (MILP).

The scope of the examined case studies is threefold. First of all, they demonstrate the interactions between the retailer, the flexible consumers and the wholesale market and analyse the fundamental effects of the consumers' time-shifting flexibility on the retailer's revenue from the consumers, its cost in the wholesale market, and its overall profit. Furthermore, the case studies analyse how these effects of demand flexibility depend on the retailer's relative size in the market and the strictness of the regulatory framework. Finally, they highlight the added value of the proposed bi-level model by comparing its outcomes against the state-of-the-art bi-level model.

3.3.1 Modeling Assumptions

For clarity reasons, the main assumptions behind the proposed model are outlined below:

1. The decision making problem of the examined retailer considers both the interaction with the served consumers (to which it sells energy) and the interaction with the wholesale market (from which it buys energy), with the overall aim of maximizing the retailer's profit.

2. The examined retailer serves a percentage β of the total demand in the system and the rest of the demand is served by other retailers and is assumed to be inflexible.

3. In contrast with the traditional fixed pricing or time-of-use pricing regimes where the offered retail prices are flat throughout the examined daily horizon or during certain intervals of this horizon (e.g. peak and off-peak periods), the examined strategic retailer can offer hour-specific retail prices to the served consumers. In order to prevent the retailer from making excessive profits at the expense of the consumers' utility, regulatory constraints are imposed on the maximum and average retail prices it can offer to its consumers [49].

4. The flexibility of served consumers is represented as a generic, technology-agnostic model [89]. This model captures two distinct aspects: their elasticity to reduce / increase their overall energy requirements and their flexibility to shift the operation of their loads in time, assuming that such shifting is energy neutral within the

examined daily horizon and does not compromise the satisfaction and comfort of the consumers.

5. Decision-making interactions between different consumers are not considered. Therefore, for computational reasons and following the similar modelling convention adopted in [26, 27, 32–36], the decision-making process of all served consumers is collectively modelled as a single optimization problem, maximizing their collective utility.

6. The wholesale market is a pool-based, energy-only market with a day-ahead horizon and hourly resolution, which is cleared by the market operator through the solution of a generation cost minimization problem. Market participants include electricity producers (submitting an increasing block-wise price-quantity offer, consisting of a number of blocks b and reflecting their cost characteristics) [89], the examined retailer and other retailers (submitting a quantity-only bid).

3.3.2 Structure of the Bi-level Optimization Model

In order to comprehensively capture the interactions between the retailer, its consumers, and the wholesale market, the proposed model is formulated as a bi-level optimization problem, the structure is illustrated in Fig. 4.1.

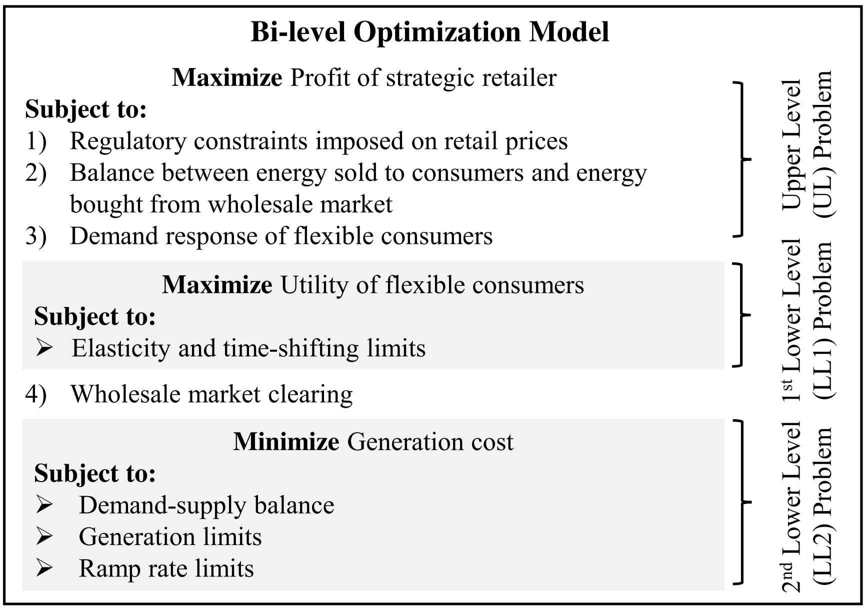

Fig. 3.1 Structure of proposed bi-level optimization model.

1. This bi-level model consists of an upper-level (UL) problem and two lower-level (LL) problems.

2. The examined retailer behaves strategically through its strategic retail pricing and bidding decisions made at the upper-level problem. The objective of the UL problem is to maximize the net profit of the retailer, and is subject to i) the regulatory constraints imposed on the offered retail prices, and ii) the balance constraint between the energy sold to the consumers and the energy bought from the wholesale market at each time period.

3. This UL problem is subject to two LL problems representing the conditions of the retail side and the wholesale side of the retailer. As shown in Fig. 4.1, those two lower-level problems respectively represent:

   (a) The demand response of retailer's served flexible consumers with the target of maximizing total demand utility and is subject to the power limits for consumption and time-coupling demand shifting constraints;

   (b) The clearing of the wholesale market with the target of minimizing the total generation cost and is subject to the demand-supply balance, the power and ramp rate limits for generation.

4. The upper-level problem and the two lower-level problems of Fig. 4.1 are all coupled as illustrated in Fig. 3.2, since:

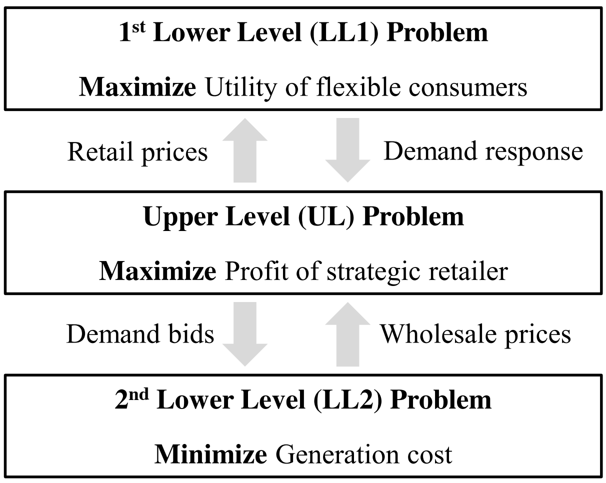

Fig. 3.2 Interrelation between the upper-level and 2 lower-level problems.

(a) The retail prices determined by the retailer (UL problem) affect the response of the consumers (as they constitute part of the objective function of LL1), while this response affects the retailer's decision making (as the consumers' demand constitutes part of the objective function and the retail balance constraint of the UL).

(b) The demand bids determined by the retailer (UL problem) affect the wholesale market clearing process (as they constitute part of the demand-supply balance constraints of LL2), while this process determines the wholesale prices which affect the retailer's decision making (as they constitute part of the objective function and average retail price constraint of the UL).

From a higher-level perspective, the proposed model captures the realistic, indirect implications of the retail prices offered by large retailers on the wholesale prices, an aspect which is not captured in the existing literature (Section 3.2). Specifically, the offered retail prices directly affect the served consumers' response; in most realistic cases where each retailer's size is substantial, this impact on the consumers' response will in turn affect the wholesale demand (as the retailer needs to buy in the wholesale market the energy demanded by the consumers) and consequently the wholesale prices.

3.4 Formulation

The proposed bi-level optimization model, the resulting MPEC and the final MILP problem are formulated in this section.

3.4.1 Decision Making of Strategic Retailer

The upper level (UL) problem represents the perspective of the examined retailer and is formulated as follows:

$$\max_{\{\lambda_t^r, d_t^w\}} \sum_t \lambda_t^r (\sum_c d_{t,c} + d_t^{sh}) - \sum_t \lambda_t^w d_t^{bid} \tag{3.1}$$

subject to:

$$\lambda^{min} \leq \lambda_t^r \leq \lambda^{max}, \forall t \tag{3.2}$$

$$\sum_t \lambda_t^r \,/\, |T| \leq K \sum_t \lambda_t^w \,/\, |T| \tag{3.3}$$

$$\sum_c d_{t,c} + d_t^{sh} = d_t^{bid}, \forall t \tag{3.4}$$

The objective function (3.1) maximizes the profit of the retailer, which is given by the difference between i) its revenue from selling energy to its consumers at the hourly retail prices λ_t^r (first term) and ii) the cost of buying energy from the wholesale market at the hourly wholesale prices λ_t^w (second term). As discussed in Section 3.2 and following the formulation presented in [37], the maximum and average retail prices are subject to the regulatory constraints (3.2) and (3.3), which aim at preventing the retailer from exploiting the consumers and making excessive profits; a stricter regulatory framework is expressed by lower values of the maximum retail price λ^{max} and the maximum ratio K between the average retail prices and the average wholesale prices. Constraint (3.4) expresses the energy balance constraints of the retailer; since the retailer acts as an intermediary between the consumers and the wholesale market and cannot physically consume or produce energy, the energy it sells to its consumers and the energy it buys from the wholesale market are equal at each hour.

3.4.2 Demand Response of Flexible Consumers

The first lower level problem (LL1) represents the collective response of all served consumers to the retail prices determined by the retailer and is formulated as follows:

$$\max_{\{d_{t,c},\, d_t^{sh}\}} \Big(\sum_{t,c} \lambda_{t,c}^D \, d_{t,c} - \sum_t \lambda_t^r \big(\sum_c d_{t,c} + d_t^{sh}\big)\Big) \tag{3.5}$$

subject to:

$$0 \leq d_{t,c} \leq d_{t,c}^{max} : \mu_{t,c}^{d-}, \mu_{t,c}^{d+}, \forall t, \forall c \tag{3.6}$$

$$\sum_t d_t^{sh} = 0 : \xi \tag{3.7}$$

$$-\alpha \sum_c d_{t,c} \leq d_t^{sh} \leq \alpha \sum_c d_{t,c} : \mu_t^{sh-}, \mu_t^{sh+}, \forall t \tag{3.8}$$

The objective function (3.5) maximizes the collective utility of the retailer's consumers, which is given by the difference between i) the benefit (or satisfaction) they perceive from the use of energy (first term) and ii) the payment to the retailer (second term). Constraints (3.6)-(3.8) express the flexibility of the consumers, which, as discussed in Section 3.2, involves both elasticity and time-shifting potentials.

The consumers' (self-price) elasticity is expressed by constraint (3.6) which implies that one part of their demand (which is denoted as baseline demand in this work) at each hour can be modified within certain limits, irrespectively of their demand levels at other hours. However, according to microeconomic principles, their perceived benefit (first term of (3.5)) is reduced with a reduction of their baseline demand. Following the approach employed in [89], their benefit function is modeled as an increasing step-wise function, consisting of a number of blocks c.

The consumers' time-shifting flexibility is expressed by constraints (3.7)-(3.8). The variable d_t^{sh} expresses the change of demand at hour t with respect to the baseline level $\sum_c d_{t,c}$ due to time-shifting, taking negative values when demand is shifted away from t and positive values when demand is shifted towards t. Constraint (3.7) represents the assumption that demand shifting is energy neutral within the daily horizon (i.e. it does not induce energy losses or gains). Constraint (3.8) expresses the limits of demand change at hour t due to time-shifting as a percentage α of the baseline level; $\alpha = 0\%$ implies that consumers do not exhibit any time-shiftable loads, while $\alpha > 0\%$ implies that a part of their energy demand can be shifted in time. The assumption that load shifting does not compromise consumers' satisfaction is expressed by the fact that their perceived benefit (first term of (3.5)) depends only on their baseline demand, while their payment to the retailer (second term of (3.5)) depends on their final demand after any potential load shifting.

3.4.3 Wholesale Market Clearing

The second lower level problem (LL2) represents the wholesale market clearing process and is formulated as:

$$\min_{\{g_{i,b,t}\}} \sum_{i,b,t} \lambda_{i,b}^{G}\, g_{i,b,t} \tag{3.9}$$

subject to:

$$d_t^{bid} + D_t^{bid} - \sum_{i,b} g_{i,b,t} = 0 : \lambda_t^{w}, \forall t \tag{3.10}$$

$$0 \leq g_{i,b,t} \leq g_{i,b}^{max} : \mu_{i,b,t}^{g-}, \mu_{i,b,t}^{g+}, \forall i, \forall b, \forall t \tag{3.11}$$

$$-R_i^D \leq \sum_b g_{i,b,t} - g_{i,0} \leq R_i^U : \mu_{i,t}^{r-}, \mu_{i,t}^{r+}, \forall i, t = 1 \tag{3.12}$$

$$-R_i^D \leq \sum_b g_{i,b,t} - \sum_b g_{i,b,(t-1)} \leq R_i^U : \mu_{i,t}^{r-}, \mu_{i,t}^{r+}, \forall i, \forall t > 1 \tag{3.13}$$

Based on the submitted retailers' bids (including the examined retailer and other retailers) and producers' offers, the market operator minimizes the total generation cost (3.9), while satisfying the market demand-supply balance constraint which ensures that the system operates in a secure fashion (3.10) (the dual variables of which constitute the wholesale prices) and the producers' operational limits (3.11) and ramp rate limits (3.12)-(3.13).

3.4.4 MPEC

In order to effectively solve the above bi-level optimization problem, the two LL problems are replaced by their respective equivalent Karush-Kuhn-Tucker (KKT) optimality conditions, a transformation enabled by the continuity and convexity of these LL problems. This converts the bi-level problem to a single-level Mathematical Program with Equilibrium Constraints (MPEC), illustrated in Fig. 3.3.

It should be noted that the optimality conditions associated with the two LL problems can be formulated through two alternative approaches: KKT conditions and primal-dual transformation. In this respect, the following two observations are relevant:

1. The first approach (KKT conditions) includes a set of complementarity conditions as a part of the optimality conditions. Such complementarity conditions can be linearized by adding a set of auxiliary binary variables.

2. The second approach (primal-dual transformation) includes the non-linear strong duality equality as a part of the optimality conditions, which can be adopted to linearize the resulting MPEC.

KKT conditions and strong duality equality associated with the two lower-level problems are derived in the next four subsections.

KKT Conditions Associated with the LL1 Problem

To obtain the KKT conditions associated with the first lower-level problem, the corresponding Lagrangian function $\mathcal{L}_{LL1}$ below is required as below:

$$\mathcal{L}_{LL1} = \sum_{t,c} \lambda_t^r d_{t,c} + \sum_t \lambda_t^r d_t^{sh} - \sum_{t,c} \lambda_{t,c}^D d_{t,c} - \sum_{t,c} \mu_{t,c}^{d-} d_{t,c} + \sum_{t,c} \mu_{t,c}^{d+} (d_{t,c} - d_{t,c}^{max}) + \\ \xi \sum_t d_t^{sh} - \sum_t \mu_t^{sh-} (d_t^{sh} + \alpha \sum_c d_{t,c}) + \sum_t \mu_t^{sh+} (d_t^{sh} - \alpha \sum_c d_{t,c}) \tag{3.14}$$

Considering the Lagrangian function $\mathcal{L}_{LL1}$ given by (3.14), the KKT first order optimality conditions of the LL1 problem are derived as follows:

$$\sum_t d_t^{sh} = 0 \tag{3.15}$$

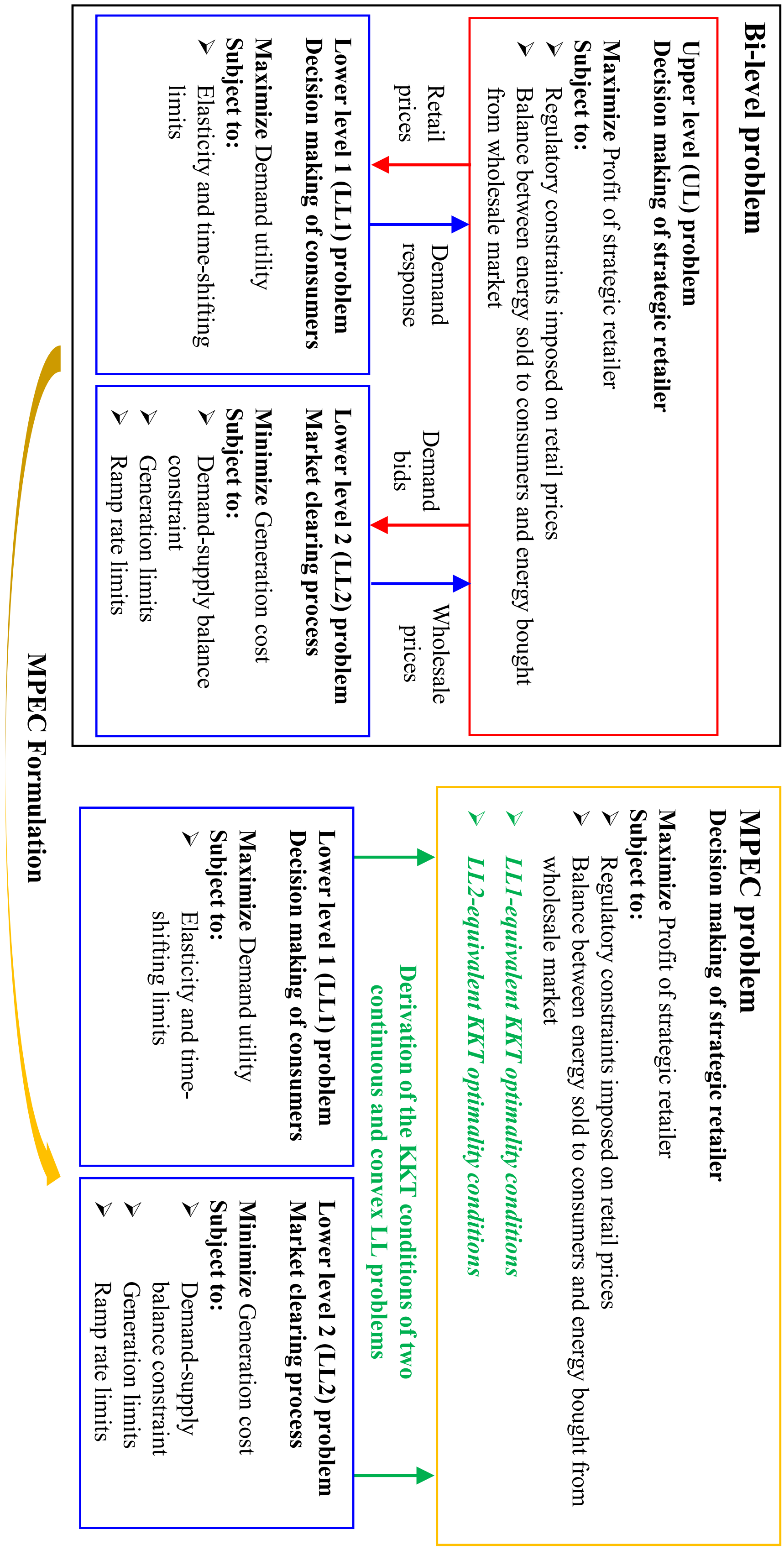

Fig. 3.3 Conversion of the bi-level model into its corresponding MPEC.

$$\frac{\partial \mathcal{L}_{LL1}}{\partial d_{t,c}} = \lambda_t^r - \lambda_{t,c}^D - \mu_{t,c}^{d-} + \mu_{t,c}^{d+} - \alpha\, \mu_t^{sh-} - \alpha\, \mu_t^{sh+} = 0, \forall t, \forall c \tag{3.16}$$

$$\frac{\partial \mathcal{L}_{LL1}}{\partial d_t^{sh}} = \lambda_t^r - \xi - \mu_t^{sh-} + \mu_t^{sh+} = 0, \forall t \tag{3.17}$$

$$0 \le \mu_{t,c}^{d-} \perp d_{t,c} \ge 0, \forall t, \forall c \tag{3.18}$$

$$0 \le \mu_{t,c}^{d+} \perp (d_{t,c}^{max} - d_{t,c}) \ge 0, \forall t, \forall c \tag{3.19}$$

$$0 \le \mu_t^{sh-} \perp (d_t^{sh} + \alpha \sum_c d_{t,c}) \ge 0, \forall t \tag{3.20}$$

$$0 \le \mu_t^{sh+} \perp (\alpha \sum_c d_{t,c} - d_t^{sh}) \ge 0, \forall t \tag{3.21}$$

The structure of the KKT conditions (3.15)-(3.21) is explained below:

1. Equality (3.15) is the primal equality constraint (3.7) in the first LL problem.
2. Equalities (3.16)-(3.17) are obtained by differentiating the Lagrangian function $\mathcal{L}_{LL1}$ with respect to the primal variables in the set of $\{d_{t,c}, d_t^{sh}\}$.
3. Complementarity conditions (3.18)-(3.21) are related to the inequality constraints (3.6) and (3.8).

KKT Conditions Associated with the LL2 Problem

To obtain the KKT conditions associated with the second LL problem, the corresponding Lagrangian function $\mathcal{L}_{LL2}$ below is required as below:

$$\begin{aligned} \mathcal{L}_{LL2} = \sum_{i,b,t} \lambda_{i,b}^G\, g_{i,b,t} + \sum_t \lambda_t^w (d_t^{bid} + D_t^{bid} - \sum_{i,b} g_{i,b,t}) - \sum_{i,b,t} \mu_{i,b,t}^{g-}\, g_{i,b,t} + \\ \sum_{i,b,t} \mu_{i,b,t}^{g+} (g_{i,b}^{max} - g_{i,b,t}) - \sum_{i,(t=1)} \mu_{i,t}^{r-} (\sum_b g_{i,b,t} - g_{i,0} + R_i^D) + \\ \sum_{i,(t=1)} \mu_{i,t}^{r+} (\sum_b g_{i,b,t} - g_{i,0} - R_i^U) - \sum_{i,(t>1)} \mu_{i,t}^{r-} (\sum_b g_{i,b,t} - \sum_b g_{i,b,(t-1)} + R_i^D) + \\ \sum_{i,(t>1)} \mu_{i,t}^{r+} (\sum_b g_{i,b,t} - \sum_b g_{i,b,(t-1)} - R_i^U) \end{aligned} \tag{3.22}$$

Considering the Lagrangian function $\mathcal{L}_{LL2}$ given by (3.22), the KKT first order optimality conditions of the first lower-level problem are derived as follows:

$$d_t^{bid} + D_t^{bid} - \sum_{i,b} g_{i,b,t} = 0, \forall t \tag{3.23}$$

$$\frac{\partial \mathcal{L}_{LL2}}{\partial g_{i,b,t}} = \lambda_{i,b}^{G} - \lambda_{t}^{w} - \mu_{i,b,t}^{g-} + \mu_{i,b,t}^{g+} - \mu_{i,t}^{r-} + \mu_{i,(t+1)}^{r-} + \mu_{i,t}^{r+} - \mu_{i,(t+1)}^{r+} = 0, \forall i, \forall b, \forall t < N_T \quad (3.24)$$

$$\frac{\partial \mathcal{L}_{LL2}}{\partial g_{i,b,t}} = \lambda_{i,b}^{G} - \lambda_{t}^{w} - \mu_{i,b,t}^{g-} + \mu_{i,b,t}^{g+} - \mu_{i,t}^{r-} + \mu_{i,t}^{r+} = 0, \forall i, \forall b, t = N_T \quad (3.25)$$

$$0 \leq \mu_{i,b,t}^{g-} \perp g_{i,b,t} \geq 0, \forall i, \forall b, \forall t \quad (3.26)$$

$$0 \leq \mu_{i,b,t}^{g+} \perp (g_{i,b}^{max} - g_{i,b,t}) \geq 0, \forall i, \forall b, \forall t \quad (3.27)$$

$$0 \leq \mu_{i,t}^{r-} \perp (\sum_{b} g_{i,b,t} - g_{i,0} + R_i^D) \geq 0, \forall i, t = 1 \quad (3.28)$$

$$0 \leq \mu_{i,t}^{r+} \perp (R_i^U - \sum_{b} g_{i,b,t} + g_{i,0}) \geq 0, \forall i, t = 1 \quad (3.29)$$

$$0 \leq \mu_{i,t}^{r-} \perp (\sum_{b} g_{i,b,t} - \sum_{b} g_{i,b,(t-1)} + R_i^D) \geq 0, \forall i, \forall t > 1 \quad (3.30)$$

$$0 \leq \mu_{i,t}^{r+} \perp (R_i^U - \sum_{b} g_{i,b,t} + \sum_{b} g_{i,b,(t-1)}) \geq 0, \forall i, \forall t > 1 \quad (3.31)$$

The structure of the KKT conditions (3.23)-(3.31) is explained below:

1. Equality (3.23) is the primal equality constraint (3.10) in the second LL problem.

2. Equality constraint (3.24)-(3.25) is obtained by differentiating the Lagrangian function $\mathcal{L}_{LL2}$ with respect to the primal variable in the set of $\{g_{i,b,t}\}$.

3. Complementarity conditions (3.26)-(3.27) are related to the inequality constraint (3.11).

4. Complementarity conditions (3.28)-(3.29) and (3.30)-(3.31) are respectively related to the inequality constraints (3.12) and (3.13).

Strong Duality Equality Associated with LL1 Problem

We consider the first LL problem is a convex optimization problem, thus we can directly get the strong duality equality through the use of primal-dual transformation:

$$\sum_{t,c} \lambda_{t,c}^{D} d_{t,c} - \sum_{t,c} \lambda_{t}^{r} d_{t,c} - \sum_{t} \lambda_{t}^{r} d_{t}^{sh} = - \sum_{t,c} \mu_{t,c}^{d+} d_{t,c}^{max} \quad (3.32)$$

where constraint (3.32) enforces the equality of the values of the primal objective function and dual objective function at the optimal solution for the 1st LL problem. This equality is used to linearize the final MPEC.

Strong Duality Equality Associated with the LL2 Problem

We consider the second LL problem is also a convex optimization problem, thus we can directly get the strong duality equality through the use of primal-dual transformation:

$$\sum_{i,b,t} \lambda^G_{i,b} g_{i,b,t} = -\sum_{i,b,t} \mu^{g+}_{i,b,t} g^{max}_{i,b} - \sum_{i,(t=1)} \mu^{r-}_{i,t} g_{i,0} + \sum_{i,(t=1)} \mu^{r+}_{i,t} g_{i,0} \tag{3.33}$$

where constraint (3.33) enforces the equality of the values of the primal objective function and dual objective function at the optimal solution for the second LL problem. This equality is used to linearize the final MPEC.

Final MPEC Formulation

A single-level MPEC corresponding to the proposed bi-level model (3.1)-(3.13) is obtained by replacing two LL problems (3.5)-(3.8) and (3.9)-(3.13) with their individual KKT conditions (3.15)-(3.21) and (3.23)-(3.31). The resulting MPEC is given below:

$$\max_{V} \sum_{t} \lambda^r_t (\sum_{c} d_{t,c} + d^{sh}_t) - \sum_{t} \lambda^w_t d^{bid}_t \tag{3.34}$$

where:

$$V = \{\lambda^r_t, d^{bid}_t, d_{t,c}, d^{sh}_t, \mu^{d-}_{t,c}, \mu^{d+}_{t,c}, \xi, \mu^{sh-}_t, \mu^{sh+}_t, g_{i,b,t}, \lambda^w_t, \mu^{g-}_{i,b,t}, \mu^{g+}_{i,b,t}, \mu^{r-}_{i,t}, \mu^{r+}_{i,t}\} \tag{3.35}$$

subject to:

$$(\text{UL constraints}) : (3.2) - (3.4) \tag{3.36}$$

$$(\text{KKT optimality conditions of the LL1 problem}) : (3.15) - (3.21) \tag{3.37}$$

$$(\text{KKT optimality conditions of the LL2 problem}) : (3.23) - (3.31) \tag{3.38}$$

The objective function of the MPEC (3.34) coincides with the objective function of the UL problem (3.1). The set of decision variables (3.35) includes i) the decision variables of the UL problem $\{\lambda^r_t, d^{bid}_t\}$; the decision variables $\{d_{t,c}, d^{sh}_t, g_{i,b,t}\}$ of two LL problems; and the Lagrangian multipliers $\{\mu^{d-}_{t,c}, \mu^{d+}_{t,c}, \xi, \mu^{sh-}_t, \mu^{sh+}_t, \lambda^w_t, \mu^{g-}_{i,b,t}, \mu^{g+}_{i,b,t}, \mu^{r-}_{i,t}, \mu^{r+}_{i,t}\}$ associated with the constraints of the two LL problems.

3.4.5 MILP

This MPEC formulation is non-linear, thus the global optimality of solutions obtained by commercial solvers is not guaranteed. Therefore, we transform this MPEC to a Mixed-Integer Linear Program (MILP) which commercial branch-and-cut solvers can efficiently solve [90].

More specifically, the above MPEC (3.34)-(3.38) includes two types of non-linearities:

1. The bilinear terms in the objective function (3.34), which is the product of energy quantity and price variables.

2. The bilinear terms in the complementarity conditions (3.18)-(3.21) and (3.26)-(3.31).

Linearizing Objective Function

Adopting the linearization approach proposed in [91], which exploits the strong duality theorem and some of the KKT equalities, the bilinear objective function is replaced with the following linear expression:
For the strong duality of the first LL problem:

$$\sum_{t,c} \lambda^D_{t,c}\, d_{t,c} - \sum_{t,c} \lambda^r_t\, d_{t,c} - \sum_{t} \lambda^r_t\, d^{sh}_t = -\sum_{t,c} \mu^{d+}_{t,c}\, d^{max}_{t,c} \tag{3.39}$$

Rearrange the equality (3.39) and the first bilinear term in the objective function (3.34):

$$\sum_{t} \lambda^r_t\, (\sum_{c} d_{t,c} + d^{sh}_t) = \sum_{t,c} \lambda^D_{t,c}\, d_{t,c} - \sum_{t,c} \mu^{d+}_{t,c}\, d^{max}_{t,c} \tag{3.40}$$

By multiplying both sides of (3.24)-(3.25) by $g_{i,b,t}$, summing for every i, b, t and rearranging some terms we get:

$$\begin{aligned}\sum_{i,b,t} \lambda^w_t\, g_{i,b,t} = \sum_{i,b,t} \lambda^G_{i,b}\, g_{i,b,t} - \sum_{i,b,t} \mu^{g-}_{i,b,t}\, g_{i,b,t} + \sum_{i,b,t} \mu^{g+}_{i,b,t}\, g_{i,b,t} - \sum_{i,b,t} \mu^{r-}_{i,t}\, g_{i,b,t} + \\ \sum_{i,b,(t<N_T)} \mu^{r-}_{i,(t+1)}\, g_{i,b,t} + \sum_{i,b,t} \mu^{r+}_{i,t}\, g_{i,b,t} - \sum_{i,b,(t<N_T)} \mu^{r+}_{i,(t+1)}\, g_{i,b,t}\end{aligned} \tag{3.41}$$

By making use of (3.4):

$$\sum_{i,b} g_{i,b,t} = d^{bid}_t + D^{bid}_t, \forall t \tag{3.42}$$

Equation (3.4.5) becomes:

$$\sum_{t} \lambda_t^w d_t^{bid} = -\sum_{t} \lambda_t^w D_t^{bid} + \sum_{i,b,t} \lambda_{i,b}^G \, g_{i,b,t} - \sum_{i,b,t} \mu_{i,b,t}^{g-} \, g_{i,b,t} + \sum_{i,b,t} \mu_{i,b,t}^{g+} \, g_{i,b,t} -$$
$$\sum_{i,b,t} \mu_{i,t}^{r-} \, g_{i,b,t} + \sum_{i,b,(t<N_T)} \mu_{i,(t+1)}^{r-} \, g_{i,b,t} + \sum_{i,b,t} \mu_{i,t}^{r+} \, g_{i,b,t} - \sum_{i,b,(t<N_T)} \mu_{i,(t+1)}^{r+} \, g_{i,b,t} \tag{3.43}$$

For complementarity conditions (3.26) and (3.27):

$$\sum_{i,b,t} \mu_{i,b,t}^{g-} \, g_{i,b,t} = 0 \tag{3.44}$$

$$\sum_{i,b,t} \mu_{i,b,t}^{g+} \, g_{i,b,t} = \sum_{i,b,t} \mu_{i,b,t}^{g+} \, g_{i,b}^{max} \tag{3.45}$$

By substituting (3.44) and (3.45) into (3.4.5), renders the equality below:

$$\sum_{t} \lambda_t^w d_t^{bid} = -\sum_{t} \lambda_t^w D_t^{bid} + \sum_{i,b,t} \lambda_{i,b}^G \, g_{i,b,t} + \sum_{i,b,t} \mu_{i,b,t}^{g+} \, g_{i,b}^{max} -$$
$$\sum_{i,b,t} \mu_{i,t}^{r-} \, g_{i,b,t} + \sum_{i,b,(t<N_T)} \mu_{i,(t+1)}^{r-} \, g_{i,b,t} + \sum_{i,b,t} \mu_{i,t}^{r+} \, g_{i,b,t} - \sum_{i,b,(t<N_T)} \mu_{i,(t+1)}^{r+} \, g_{i,b,t} \tag{3.46}$$

For complementarity conditions (3.28)-(3.31):

$$\sum_{i,b,t} \mu_{i,t}^{r-} \, g_{i,b,t} - \sum_{i,b,(t=1)} \mu_{i,t}^{r-} \, g_{i,0} - \sum_{i,b,(t>1)} \mu_{i,t}^{r-} \, g_{i,b,(t-1)} = -\sum_{i,b,t} \mu_{i,t}^{r-} \, R_i^D \tag{3.47}$$

$$\sum_{i,b,t} \mu_{i,t}^{r+} \, g_{i,b,t} - \sum_{i,b,(t=1)} \mu_{i,t}^{r+} \, g_{i,0} - \sum_{i,b,(t>1)} \mu_{i,t}^{r+} \, g_{i,b,(t-1)} = \sum_{i,b,t} \mu_{i,t}^{r+} \, R_i^U \tag{3.48}$$

Renders the equalities (3.47) and (3.48) below:

$$\sum_{i,b,t} \mu_{i,t}^{r-} \, g_{i,b,t} - \sum_{i,b,(t<N_T)} \mu_{i,(t+1)}^{r-} \, g_{i,b,t} = -\sum_{i,b,t} \mu_{i,t}^{r-} \, R_i^D + \sum_{i,b,(t=1)} \mu_{i,t}^{r-} \, g_{i,0} \tag{3.49}$$

$$\sum_{i,b,t} \mu_{i,t}^{r+} \, g_{i,b,t} - \sum_{i,b,(t<N_T)} \mu_{i,(t+1)}^{r+} \, g_{i,b,t} = \sum_{i,b,t} \mu_{i,t}^{r+} \, R_i^U + \sum_{i,b,(t=1)} \mu_{i,t}^{r+} \, g_{i,0} \tag{3.50}$$

By substituting (3.49) and (3.50) into (3.4.5), renders the equality below:

$$\sum_{t} \lambda_t^w d_t^{bid} = -\sum_{t} \lambda_t^w D_t^{bid} + \sum_{i,b,t} \lambda_{i,b}^G \, g_{i,b,t} + \sum_{i,b,t} \mu_{i,b,t}^{g+} \, g_{i,b}^{max} +$$
$$\sum_{i,b,t} \mu_{i,t}^{r-} \, R_i^D - \sum_{i,b,(t=1)} \mu_{i,t}^{r-} \, g_{i,0} + \sum_{i,b,t} \mu_{i,t}^{r+} \, R_i^U + \sum_{i,b,(t=1)} \mu_{i,t}^{r+} \, g_{i,0} \tag{3.51}$$

Therefore, the bilinear terms $\sum_t \lambda_t^r (\sum_c d_{t,c} + d_t^{sh}) - \sum_t \lambda_t^w d_t^{bid}$ in the objective function (3.34) of the MPEC problem can be replaced with the expressions in the right side of (3.40) and (3.4.5) which are linear, and the objective function of the MILP problem is:

$$\sum_{t,c} \lambda_{t,c}^D \, d_{t,c} - \sum_{t,c} \mu_{t,c}^{d+} \, d_{t,c}^{max} + \sum_t \lambda_t^w D_t^{bid} - \sum_{i,b,t} \lambda_{i,b}^G \, g_{i,b,t} - \sum_{i,b,t} \mu_{i,b,t}^{g+} \, g_{i,b}^{max} - \\ \sum_{i,b,t} \mu_{i,t}^{r-} \, R_i^D + \sum_{i,b,(t=1)} \mu_{i,t}^{r-} \, g_{i,0} - \sum_{i,b,t} \mu_{i,t}^{r+} \, R_i^U - \sum_{i,b,(t=1)} \mu_{i,t}^{r+} \, g_{i,0} \tag{3.52}$$

Linearizing complementarity conditions

The bilinear terms in the complementarity conditions (3.18)-(3.21) and (3.26)-(3.31) can be expressed in the generic form $0 \leq \mu \perp p \geq 0$, with μ and p representing generic dual and primal terms respectively. The linearization approach proposed in [92] replaces each of these conditions with the set of mixed-integer linear conditions $\mu \geq 0$, $p \geq 0$, $\mu \leq \omega M^\mu$, $p \leq (1-\omega) M^p$, where ω is an auxiliary binary variable, while M^μ and M^p are large positive constants.

The values of the parameters M^D and M^P should be suitably selected in order to achieve not only accurate but also computationally efficient solution of the MILP. Specifically, M^D and M^P should be large enough in order to avoid imposing additional upper bounds on the decision variables and thus resulting in an inaccurate solution of the MILP. On the other hand, extremely large values should be avoided as they hinder the convergence of branch-and-cut solvers and result in large computational times [91]. Suitable values of the parameters M^P corresponding to primal terms can be more easily determined based on the bounds of primal variables which correspond to explicit physical limits. For example, the parameter M^P corresponding to the primal term of the complementarity constraint (3.18) is set equal to the maximum demand limit $d_{t,c}^{max}$ which physically limits the primal variable $d_{t,c}$. Suitable selection of the parameters M^D corresponding to dual terms is more challenging since the dual variables do not exhibit explicit physical limits. In this context, the heuristic approach presented in [91] has been employed to tune parameters M^D.

Final MILP Formulation

Considering the linearization techniques presented above, MPEC (3.34)-(3.38) can be transformed into the MILP problem given by (3.4.5)-(3.112). Where the set of decision variables of the MILP formulation includes the set (3.54) as well as the auxiliary binary variables (3.55) introduced for linearizing (3.18)-(3.21) and (3.26)-(3.31).

$$\max_{V,S} \sum_{t,c} \lambda^{D}_{t,c}\, d_{t,c} - \sum_{t,c} \mu^{d+}_{t,c}\, d^{max}_{t,c} + \sum_{t} \lambda^{w}_{t}\, D^{bid}_{t} - \sum_{i,b,t} \lambda^{G}_{i,b}\, g_{i,b,t} - \sum_{i,b,t} \mu^{g+}_{i,b,t}\, g^{max}_{i,b} - \sum_{i,b,t} \mu^{r-}_{i,t}\, R^{D}_{i} + \sum_{i,b,(t=1)} \mu^{r-}_{i,t}\, g_{i,0} - \sum_{i,b,t} \mu^{r+}_{i,t}\, R^{U}_{i} - \sum_{i,b,(t=1)} \mu^{r+}_{i,t}\, g_{i,0} \tag{3.53}$$

where:

$$V = \{\lambda^{r}_{t}, d^{bid}_{t}, d_{t,c}, d^{sh}_{t}, \mu^{d-}_{t,c}, \mu^{d+}_{t,c}, \xi, \mu^{sh-}_{t}, \mu^{sh+}_{t}, g_{i,b,t}, \lambda^{w}_{t}, \mu^{g-}_{i,b,t}, \mu^{g+}_{i,b,t}, \mu^{r-}_{i,t}, \mu^{r+}_{i,t}\} \tag{3.54}$$

$$S = \{\omega^{d-}_{t,c}, \omega^{d+}_{t,c}, \omega^{sh-}_{t}, \omega^{sh+}_{t}, \omega^{g-}_{i,b,t}, \omega^{g+}_{i,b,t}, \omega^{r-}_{i,t}, \omega^{r+}_{i,t}\} \tag{3.55}$$

subject to:
UL constraints:

$$\lambda^{min} \leq \lambda^{r}_{t} \leq \lambda^{max}, \forall t \tag{3.56}$$

$$\sum_{t} \lambda^{r}_{t} \,/\, |T| \leq K \sum_{t} \lambda^{w}_{t} \,/\, |T| \tag{3.57}$$

$$\sum_{c} d_{t,c} + d^{sh}_{t} = d^{bid}_{t}, \forall t. \tag{3.58}$$

KKT optimality conditions and linearized complementary conditions of the 1st LL problem:

$$\sum_{t} d^{sh}_{t} = 0 \tag{3.59}$$

$$\lambda^{r}_{t} - \lambda^{D}_{t,c} - \mu^{d-}_{t,c} + \mu^{d+}_{t,c} - \alpha\, \mu^{sh-}_{t} - \alpha\, \mu^{sh+}_{t} = 0, \forall t, \forall c \tag{3.60}$$

$$\lambda^{r}_{t} - \xi - \mu^{sh-}_{t} + \mu^{sh+}_{t} = 0, \forall t \tag{3.61}$$

$$\mu^{d-}_{t,c} \geq 0, \forall t, \forall c \tag{3.62}$$

$$d_{t,c} \geq 0, \forall t, \forall c \tag{3.63}$$

$$\mu^{d-}_{t,c} \leq \omega^{d-}_{t,c}\, M^{\mu}, \forall t, \forall c \tag{3.64}$$

$$d_{t,c} \leq (1 - \omega^{d-}_{t,c})\, M^{P}, \forall t, \forall c \tag{3.65}$$

$$\mu^{d+}_{t,c} \geq 0, \forall t, \forall c \tag{3.66}$$

$$d^{max}_{t,c} - d_{t,c} \geq 0, \forall t, \forall c \tag{3.67}$$

$$\mu^{d+}_{t,c} \leq \omega^{d+}_{t,c}\, M^{\mu}, \forall t, \forall c \tag{3.68}$$

$$d^{max}_{t,c} - d_{t,c} \leq (1 - \omega^{d+}_{t,c})\, M^{P}, \forall t, \forall c \tag{3.69}$$

$$\mu_t^{sh-} \geq 0, \forall t \tag{3.70}$$

$$d_t^{sh} + \alpha \sum_c d_{t,c} \geq 0, \forall t \tag{3.71}$$

$$\mu_{t,c}^{sh-} \leq \omega_{t,c}^{sh-} M^{\mu}, \forall t \tag{3.72}$$

$$d_t^{sh} + \alpha \sum_c d_{t,c} \leq (1 - \omega_{t,c}^{sh-})\, M^P, \forall t \tag{3.73}$$

$$\mu_t^{sh+} \geq 0, \forall t \tag{3.74}$$

$$d_t^{sh} - \alpha \sum_c d_{t,c} \geq 0, \forall t \tag{3.75}$$

$$\mu_{t,c}^{sh+} \leq \omega_{t,c}^{sh+} M^{\mu}, \forall t \tag{3.76}$$

$$d_t^{sh} - \alpha \sum_c d_{t,c} \leq (1 - \omega_{t,c}^{sh+})\, M^P, \forall t \tag{3.77}$$

$$\omega_{t,c}^{d-} \in \{0,1\}, \forall t, \forall c \tag{3.78}$$

$$\omega_{t,c}^{d+} \in \{0,1\}, \forall t, \forall c \tag{3.79}$$

$$\omega_t^{sh-} \in \{0,1\}, \forall t \tag{3.80}$$

$$\omega_t^{sh+} \in \{0,1\}, \forall t \tag{3.81}$$

KKT optimality conditions and linearized complementary conditions of the 2nd LL problem:

$$d_t^{bid} + D_t^{bid} - \sum_{i,b} g_{i,b,t} = 0, \forall t \tag{3.82}$$

$$\lambda_{i,b}^G - \lambda_t^w - \mu_{i,b,t}^{g-} + \mu_{i,b,t}^{g+} - \mu_{i,t}^{r-} + \mu_{i,(t+1)}^{r-} + \mu_{i,t}^{r+} - \mu_{i,(t+1)}^{r+} = 0, \forall i, \forall b, \forall t < N_T \tag{3.83}$$

$$\lambda_{i,b}^G - \lambda_t^w - \mu_{i,b,t}^{g-} + \mu_{i,b,t}^{g+} - \mu_{i,t}^{r-} + \mu_{i,t}^{r+} = 0, \forall i, \forall b, t = N_T \tag{3.84}$$

$$\mu_{i,b,t}^{g-} \geq 0, \forall i, \forall b, \forall t \tag{3.85}$$

$$g_{i,b,t} \geq 0, \forall i, \forall b, \forall t \tag{3.86}$$

$$\mu_{i,b,t}^{g-} \leq \omega_{i,b,t}^{g-} M^{\mu}, \forall i, \forall b, \forall t \tag{3.87}$$

$$g_{i,b,t} \leq (1 - \omega_{i,b,t}^{g-})\, M^P, \forall i, \forall b, \forall t \tag{3.88}$$

$$\mu_{i,b,t}^{g+} \geq 0, \forall i, \forall b, \forall t \tag{3.89}$$

$$g_{i,b}^{max} - g_{i,b,t} \geq 0, \forall i, \forall b, \forall t \tag{3.90}$$

$$\mu_{i,b,t}^{g+} \leq \omega_{i,b,t}^{g+} M^{\mu}, \forall i, \forall b, \forall t \tag{3.91}$$

$$g_{i,b}^{max} - g_{i,b,t} \leq (1 - \omega_{i,b,t}^{g+})\, M^P, \forall i, \forall b, \forall t \tag{3.92}$$

$$\mu_{i,t}^{r-} \geq 0, \forall i, t = 1 \tag{3.93}$$

$$\sum_b g_{i,b,t} - g_{i,0} + R_i^D \geq 0, \forall i, t = 1 \tag{3.94}$$

$$\mu_{i,t}^{r-} \leq \omega_{i,t}^{r-}\, M^\mu, \forall i, t = 1 \tag{3.95}$$

$$\sum_b g_{i,b,t} - g_{i,0} + R_i^D \leq (1 - \omega_{i,t}^{r-})\, M^P, \forall i, t = 1 \tag{3.96}$$

$$\mu_{i,t}^{r+} \geq 0, \forall i, t = 1 \tag{3.97}$$

$$R_i^U - \sum_b g_{i,b,t} + g_{i,0} \geq 0, \forall i, t = 1 \tag{3.98}$$

$$\mu_{i,t}^{r+} \leq \omega_{i,t}^{r+}\, M^\mu, \forall i, t = 1 \tag{3.99}$$

$$R_i^U - \sum_b g_{i,b,t} + g_{i,0} \leq (1 - \omega_{i,t}^{r+})\, M^P, \forall i, t = 1 \tag{3.100}$$

$$\mu_{i,t}^{r-} \geq 0, \forall i, \forall t > 1 \tag{3.101}$$

$$\sum_b g_{i,b,t} - \sum_b g_{i,b,(t-1)} + R_i^D \geq 0, \forall i, \forall t > 1 \tag{3.102}$$

$$\mu_{i,t}^{r-} \leq \omega_{i,t}^{r-}\, M^\mu, \forall i, \forall t > 1 \tag{3.103}$$

$$\sum_b g_{i,b,t} - \sum_b g_{i,b,(t-1)} + R_i^D \leq (1 - \omega_{i,t}^{r-})\, M^P, \forall i, \forall t > 1 \tag{3.104}$$

$$\mu_{i,t}^{r+} \geq 0, \forall i, \forall t > 1 \tag{3.105}$$

$$R_i^U - \sum_b g_{i,b,t} + \sum_b g_{i,b,(t-1)} \geq 0, \forall i, \forall t > 1 \tag{3.106}$$

$$\mu_{i,t}^{r+} \leq \omega_{i,t}^{r+}\, M^\mu, \forall i, \forall t > 1 \tag{3.107}$$

$$R_i^U - \sum_b g_{i,b,t} + \sum_b g_{i,b,(t-1)} \leq (1 - \omega_{i,t}^{r+})\, M^P, \forall i, \forall t > 1 \tag{3.108}$$

$$\omega_{i,b,t}^{g-} \in \{0, 1\}, \forall i, \forall b, \forall t \tag{3.109}$$

$$\omega_{i,b,t}^{g+} \in \{0, 1\}, \forall i, \forall b, \forall t \tag{3.110}$$

$$\omega_{i,t}^{r-} \in \{0, 1\}, \forall i, \forall t \tag{3.111}$$

$$\omega_{i,t}^{r+} \in \{0, 1\}, \forall i, \forall t \tag{3.112}$$

3.5 Case Studies

3.5.1 Test Data and Implementation

By employing the model presented in Section 3.4, the examined case studies aim at quantitatively demonstrating and analyzing the effects of demand flexibility on the business case of an examined electricity retailer, including its revenue from selling energy to its served consumers, its cost for buying energy from the wholesale market and its profit. As discussed in Section 3.3.1, the proposed model captures both self-price elasticity and time-shifting aspects of demand flexibility. However, the self-price elasticity is generally very low, and the time-shifting flexibility is regarded as the most promising flexibility potential, since consumers are more likely to shift their demand towards preceding or succeeding periods rather than completely curtailing it [2]. In this context, the case studies examine different scenarios regarding the extent of this time-shifting flexibility, as expressed by parameter α; the scenario without time-shifting flexibility ($\alpha = 0\%$) is referred as the benchmark scenario in the remainder.

Furthermore, as discussed in Sections 3.2 and 3.3, in contrast with the existing literature, the proposed model can capture the impact of the retail prices offered by the examined retailer on the wholesale prices, which becomes particularly important as the size of the retailer increases. In this context, the case studies also examine different scenarios regarding the relative size of the examined retailer, as expressed by parameter β.

Finally, given that the retailer's business case depends on potential regulatory constraints discussed in Section 3.3, the case studies examine different scenarios regarding the maximum allowable retail price λ^{max} (while the minimum allowable retail price $\lambda^{min} = 0$ and the maximum ratio K between the average retail prices and the average wholesale prices is assumed equal to 1 in all examined studies).

The case studies apply the proposed model in the context of a single day. The examined wholesale market reflects the general properties of the GB power system and includes 7 electricity producers, the cost parameters and maximum output limits of which are given in Table 3.2. The system demand profile D_t of the single day is presented in Fig. 3.4, while the maximum demand of the consumers served by the examined retailer is calculated by the parameter β, i.e., $\sum_c d_{t,c}^{max} = \beta \times D_t$ and the demand bid by other retailers in the wholesale market is expressed as $D_t^{bid} = (1-\beta) \times D_t$. The hourly values of the linear benefit coefficient of the consumers are illustrated in Fig. 3.5.

Table 3.2 Operational parameters of 7 electricity producers.

Generation company i	1	2	3	4	5	6	7
l_i^G (£/MWh)	10	15	23	35	50	70	100
q_i^G (£/MW^{2}h)	0.0001	0.0006	0.0014	0.0026	0.0042	0.0065	0.001
g_i^{max} (MW)	13,170	11,520	7,560	6,670	6,500	5,760	5,500

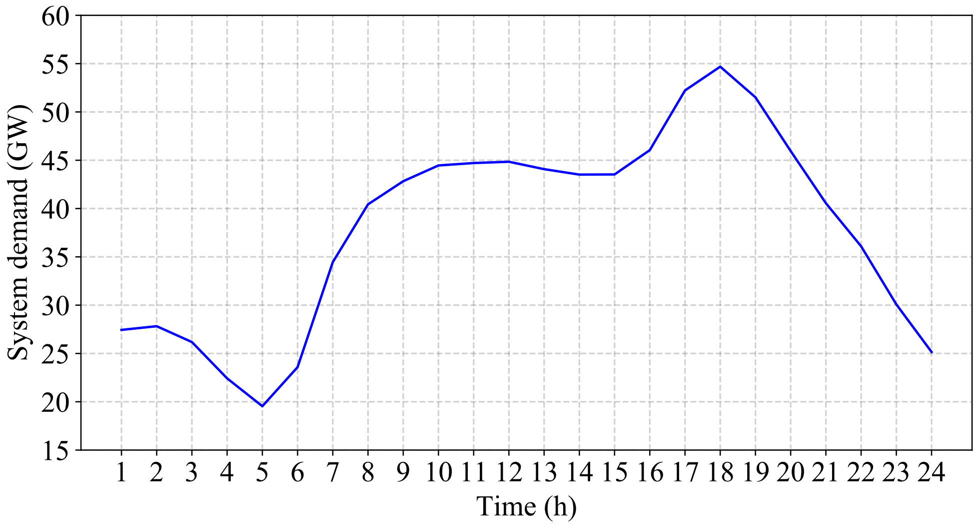

Fig. 3.4 System demand profile of the examined day.

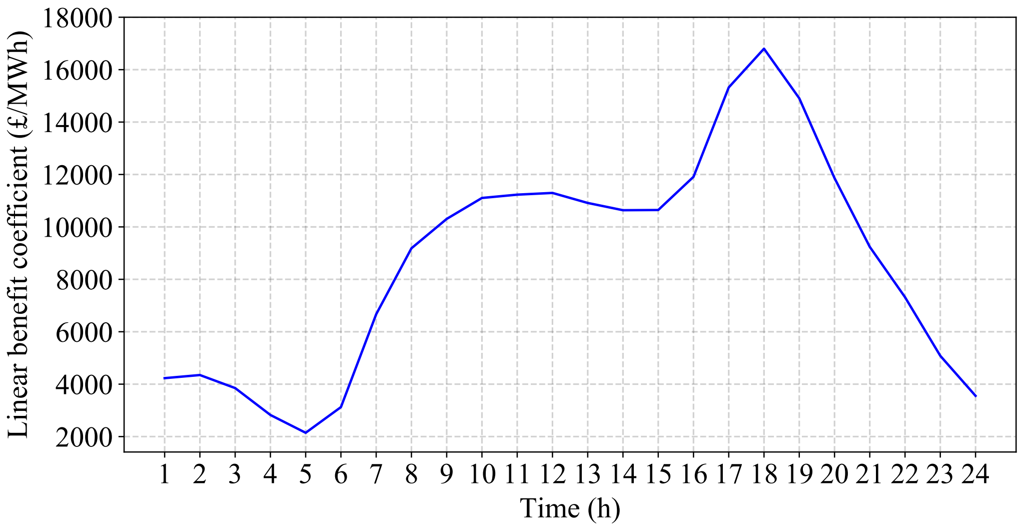

Fig. 3.5 Hourly values of linear benefit coefficient of consumers.

The final MILP model has been implemented and solved using the optimization software FICOTM Xpress [93] on a computer with a 6-core 3.50 GHz Intel(R) Xeon(R) E5-1650 v3 processor and 32 GB of RAM. The total number of binary decision variables, continuous decision variables and constraints of this MILP is 2304, 3361 and 10298, respectively. The average computational time required for solving this MILP (within a mixed-integer programming (MIP) gap lower than 0.01%) across all the examined scenarios was around 1.7s.

3.5.2 Impact of Demand Flexibility

The aim of the first set of studies lies in analyzing the fundamental effects of demand flexibility on the business case of the examined retailer, by executing the proposed model for different values of the time-shifting flexibility α of the retailer's consumers, while assuming that the relative size of the retailer is $\beta = 30\%$, and the maximum retail price is $\lambda^{max} = 200$£/MWh.

As discussed in Section 3.3, the first interaction captured by the proposed model is the one between the retail prices offered by the retailer and its consumers' demand response. Fig. 3.6 and 3.7 demonstrate this interaction by illustrating the hourly demand of the consumers (after any potential load shifting, i.e. $\sum_c d_{t,c} + d_t^{sh}$) and the hourly retail prices offered by the retailer (λ_t^r) in the examined day, for different time-shifting flexibility scenarios.

When the consumers do not exhibit any time-shifting flexibility ($\alpha = 0$), the retailer offers the highest allowable price ($\lambda^{max} = 200$£/MWhh) at peak hours (periods with the highest demand) and the lowest price (0) at off-peak hours (periods with the lowest demand) (Fig. 3.7), in order to maximize its revenue, which is determined by the summation of the demand-price products across all hours, but also satisfy the regulatory constraint imposed on the average retail price (3.3).

However, when the consumers exhibit some time-shifting flexibility ($\alpha > 0$), they are able to respond to the hour-specific prices offered by the retailer and shift a part of their demand from high-price hours to low-price hours in order to maximize their utility, with this shifting effect being enhanced with a higher α (Fig. 3.6). Since the strategic retailer anticipates this shifting response by its consumers (by employing the proposed model), it offers higher prices at off-peak hours and lower prices at peak hours, in order to maintain its revenue at the highest possible level. Consequently, a higher time-shifting flexibility results in a flatter retail price profile (Fig. 3.7).

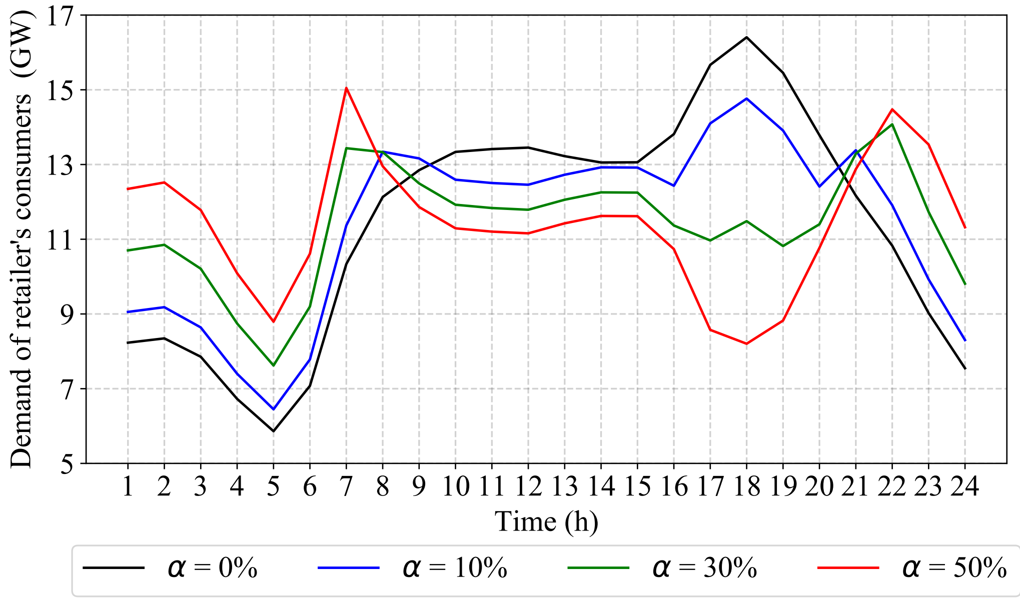

Fig. 3.6 Hourly demand of retailer's consumers for different demand flexibility scenarios.

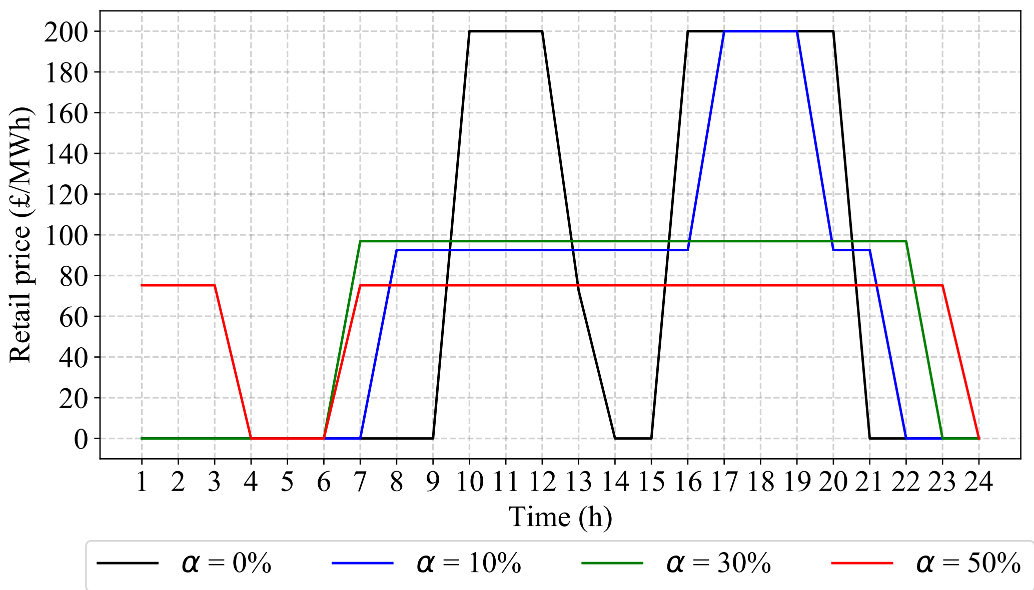

Fig. 3.7 Hourly retail prices for different demand flexibility scenarios.

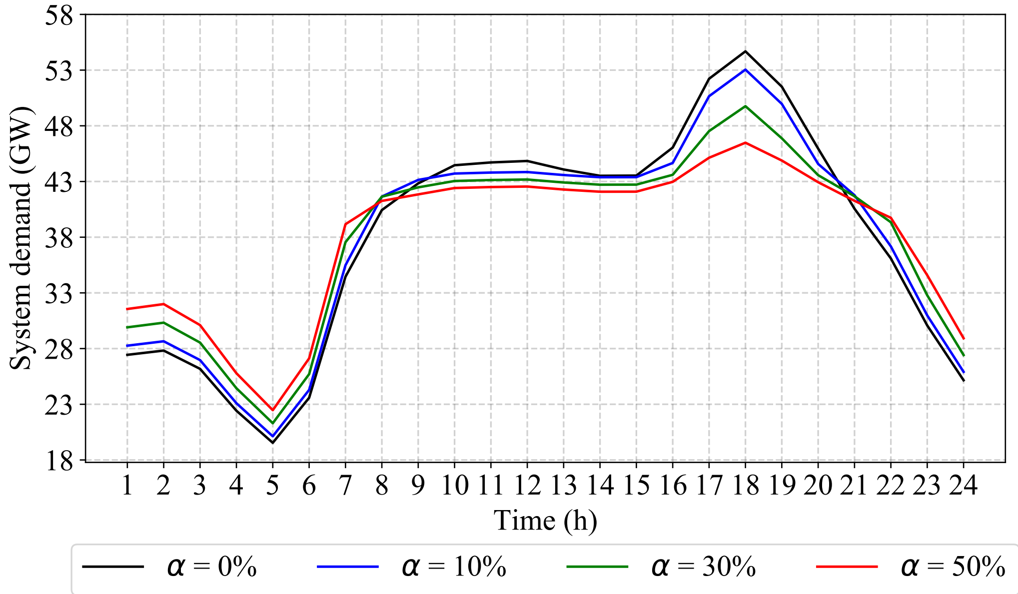

Fig. 3.8 Hourly system demand for different demand flexibility scenarios.

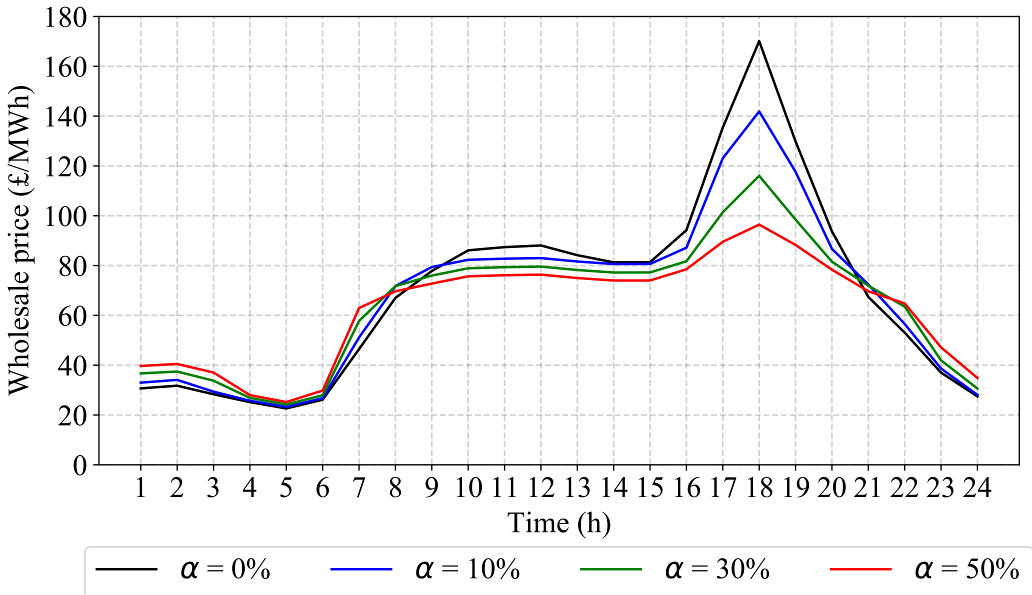

Fig. 3.9 Hourly wholesale market prices for different demand flexibility scenarios.

The second interaction captured by the proposed model is the one between the decision making of the retailer and the wholesale market. Fig. 3.8 and 3.9 demonstrate this interaction by illustrating the hourly system demand ($d_t^{bid} + D_t^{bid}$) and the hourly wholesale market prices (λ_t^w) in examined day, for different demand flexibility scenarios.

Since the retailer needs to buy from the wholesale market the energy sold to its consumers at each hour, and given that its share in the market is substantial ($\beta = 30\%$), the demand shifting response of its consumers drives a flattening effect on the system demand and subsequently on the wholesale market prices, which is enhanced with a higher α (Fig. 3.8 and 3.9). Furthermore, it should be noted that although the demand flattening effect is balanced i.e. the overall demand reduction during peak hours is equal to the overall demand increase during off-peak hours (due to the assumed energy neutrality of demand shifting (3.7)), the same does not apply to the wholesale price flattening effect. Specifically, the wholesale price reduction during peak hours is significantly more substantial than the wholesale price increase during off-peak hours. This trend is driven by the increasing slope of the producers' marginal cost curve as the demand level increases [2]. As a result of this trend, the average wholesale price over the examined day is reduced and subsequently, given the regulatory constraint imposed the average retail price (3.3) and the effort of the retailer to maximize its profit, the average retail price is accordingly reduced (Table 3.3).

In line with the discussion in Section 3.2, it should be stressed at this point that the above effects on the wholesale prices (flattened profile and reduced average) cannot be captured by the models proposed in the existing literature, as they treat the wholesale prices as exogenous, fixed that are not affected by the retailer's pricing decisions and the subsequent response of its consumers. The proposed model is able to capture these effects by representing endogenously the wholesale market clearing process.

Table 3.3 Average retail and wholesale price for different demand flexibility scenarios.

Demand Flexibility	$\alpha = 0\%$	$\alpha = 10\%$	$\alpha = 30\%$	$\alpha = 50\%$
Average Price (£/MWh)	69.70	67.40	64.59	62.69

Having demonstrated the interactions between the retailer, its consumers' response and the wholesale market, the final part of this analysis lies in investigating what are the overall effects of demand flexibility on the business case of the examined retailer. Fig. 3.10 illustrates the overall revenue, cost and profit of the retailer over the examined day, for different time-shifting flexibility scenarios.

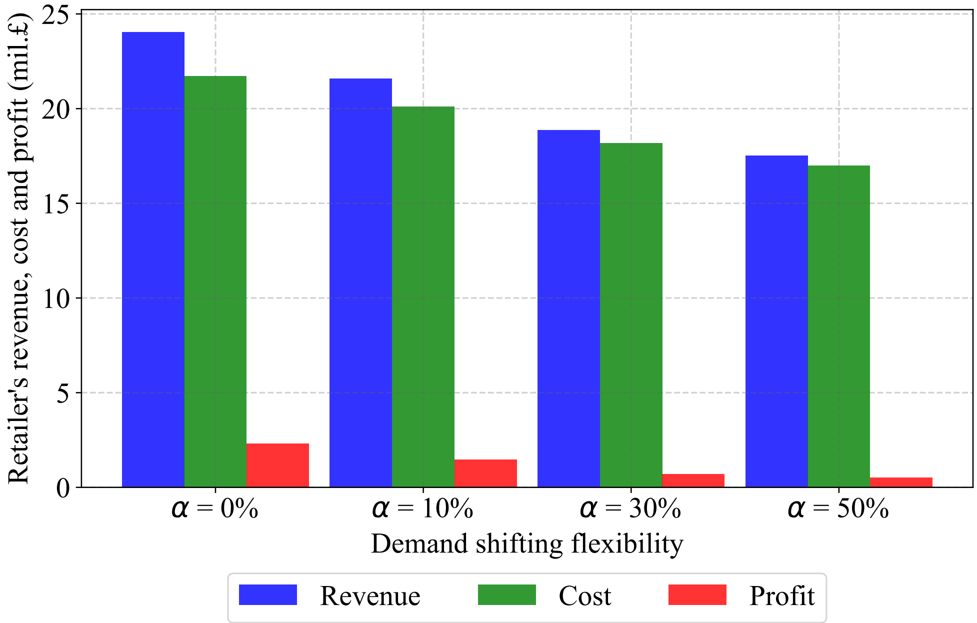

Fig. 3.10 Retailer's revenue, cost and profit for different demand flexibility scenarios.

First of all, the total revenue of the retailer is reduced with a higher α. This trend is driven by the following effects of demand flexibility: i) the consumers' demand and the offered retail prices become flatter (Fig. 3.6 and 3.7), and the retailer's revenue is given by the summation of the demand-retail price products across all hours, and ii) the average retail price is reduced (Table 3.3); it should be noted that the first effect has a more prominent role in the reduction of the retailer's revenue. This trend implies that demand flexibility deteriorates the business case of the retailer, since it limits the strategic potential of exploiting the consumers through setting large retail price differentials between peak and off-peak hours (Fig. 3.7).

Furthermore, the total cost of the retailer is also reduced with a higher α. This trend is driven by the following effects of demand flexibility: i) the consumers' demand and the wholesale prices become flatter (Fig. 3.6 and 3.9), and the retailer's cost is given by the summation of the demand-wholesale price products across all hours, and ii) the average wholesale price is reduced (Table 3.3); the first effect has a more prominent role in the reduction of the retailer's cost. This trend implies that demand flexibility improves the business case of the retailer, since it enables the retailer to a) buy more energy at low-price (off-peak) hours but also b) reduce the wholesale prices at high-price (peak) hours.

Finally, the effect of demand flexibility on the total profit of the retailer (which constitutes the most important index of the retailer's business case) is naturally driven by the combination of its effects on the retailer's revenue and cost. Since these effects counteract each other, i.e. demand flexibility reduces its profit by reducing its revenue but also increases its profit by reducing its cost, the overall effect on the retailer's profit depends on which of these two profit components is reduced at a higher rate. For the scenario investigated so far, the revenue exhibits a higher rate of reduction than the cost and consequently the profit is reduced with higher demand flexibility (Fig. 3.10). However, this trend is not definite and it can be reversed, depending on other parameters of the problem, analyzed in the later Sections 3.5.3 and 3.5.4. After analyzing how the demand flexibility affects the retailer's business, it would be also interesting to benchmark the simulated results against a perfectly coordinated wholesale / retail market without retailer interaction, that would expect a lower energy cost in both retail and wholesale market.

3.5.3 Impact of Market Competition

As discussed above in Section 3.5.2, the profit trend of the examined retailer is related to the parameters of the problem. This section aims at analyzing how these effects depend on one such parameter, the strictness of the imposed regulatory constraints on the retail prices. In this context and as discussed in Section 3.5.1, we examine different scenarios regarding the maximum allowable retail price λ^{max} for a certain extent of time-shifting flexibility ($\alpha = 30\%$) and the same relative size of the retailer ($\beta = 30\%$).

Fig. 3.11 illustrates the hourly retail prices offered by the retailer in the examined day, for the examined time-shifting flexibility scenario and the benchmark scenario without demand flexibility ($\alpha = 0\%$), and three different scenarios regarding λ^{max}. As previously discussed, when the consumers do not exhibit flexibility, the retailer offers λ^{max} at peak hours; as a result, a stricter regulatory constraint on λ^{max} reduces the revenue of the retailer. However, when the consumers exhibit flexibility of $\alpha = 30\%$, the offered retail prices are flatter and do not reach λ^{max} at any hour, irrespectively of the value of λ^{max}. In other words, the strictness of the regulatory constraint on λ^{max} does not affect the offered retail prices and the revenue of the retailer in this case.

Fig. 3.12 demonstrates the effect of demand flexibility on the business case of the examined retailer, by illustrating the reduction of its revenue, cost and profit driven by the introduction of time-shifting flexibility $\alpha = 30\%$ with respect to the benchmark scenario ($\alpha = 0\%$), for the different scenarios regarding λ^{max}. Following the above

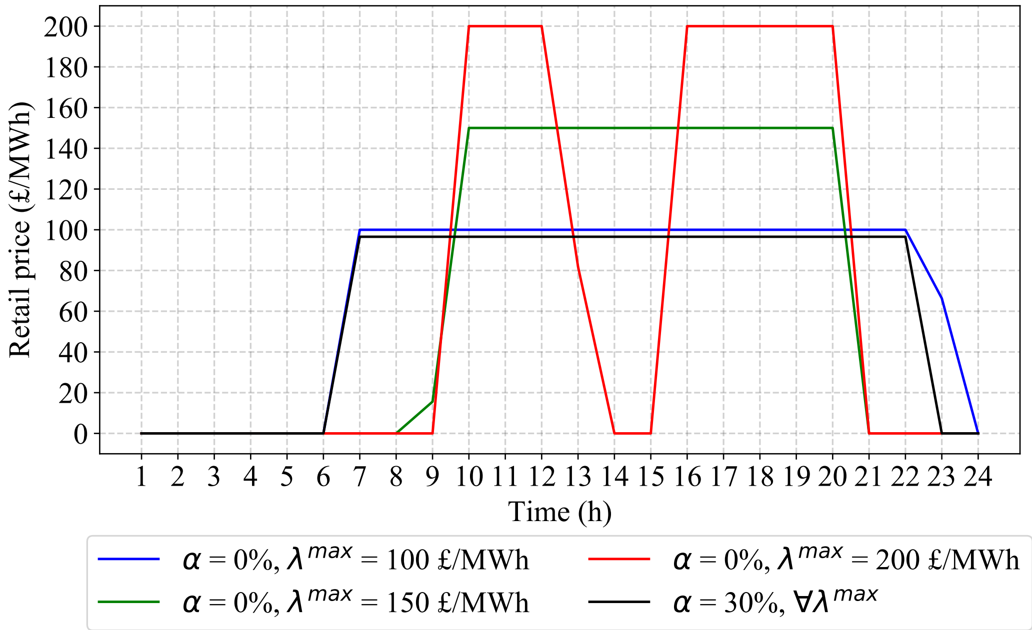

Fig. 3.11 Hourly retail prices for different demand flexibility and maximum retail price scenarios.

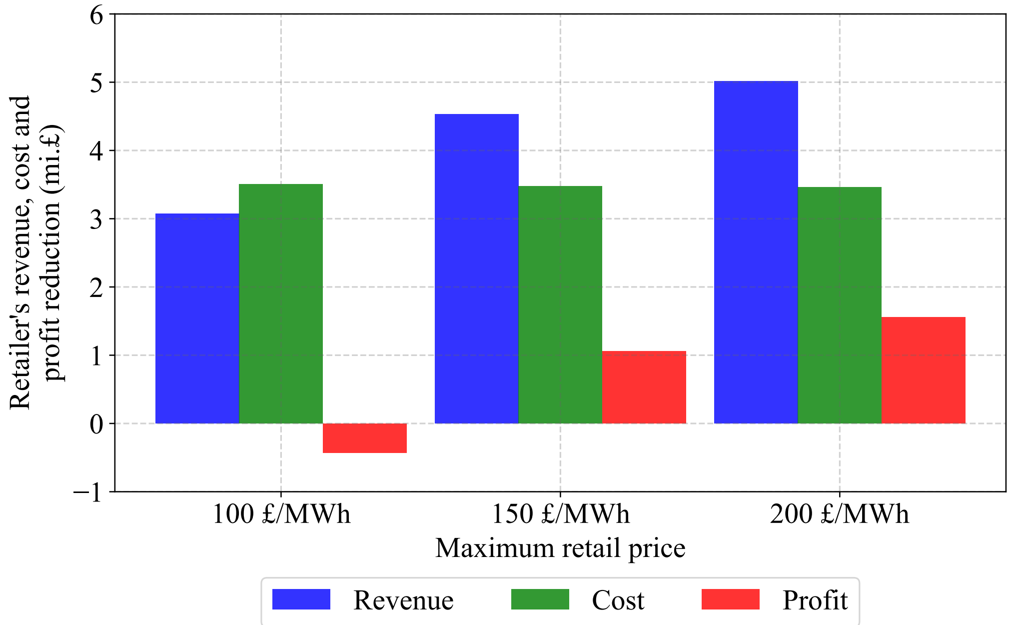

Fig. 3.12 Retailer's revenue, cost and profit reduction brought by demand flexibility of $\alpha = 30\%$ for different maximum retail price scenarios.

discussion concerning the impact of λ^{max} under $\alpha = 0\%$ and $\alpha = 30\%$, the deteriorating effect of demand flexibility on the retailer's revenue is diminished as λ^{max} reduces. On the other hand, λ^{max} does not affect that wholesale market conditions and does not alter the positive effect of demand flexibility on the retailer's cost. As a result, for the lowest value of the maximum retail price ($\lambda^{max} = 100$£/MWh), the revenue exhibits a lower reduction than the cost and consequently demand flexibility increases the retailer's profit (as demonstrated by the negative reduction in Fig. 3.12). This is the reverse trend with respect to scenarios with higher λ^{max}.

These results imply that a stricter regulatory framework limits the retailer's strategic potential of exploiting inflexible consumers through setting large price differentials between peak and off-peak hours, and therefore the impact of demand flexibility in limiting this potential is diminished. Therefore, in this case, demand flexibility improves the overall business case of the retailer since its most important impact is the improvement of the retailer's position in the wholesale market. On the other hand, under a looser regulatory framework, demand flexibility deteriorates the business case of the retailer since its most important impact is the limitation of the retailer's strategic potential.

3.5.4 Impact of Retailer's Relative Size

After having analyzed the fundamental effects of demand flexibility on the retailer's business case, the aim of this section lies in analyzing how these effects depend on the retailer's relative size. As discussed in Section 3.1, this analysis is highly relevant for this work, as the ability of the proposed model to capture the impact of the retailer's pricing strategies on the wholesale market becomes particularly important as the size of the retailer increases. In this context and in contrast with the previous Section, the proposed model is executed for different values of the relative size β of the examined retailer, while assuming that the maximum retail price is $\lambda^{max} = 200$£/MWhh.

Fig. 3.13, 3.14 and 3.15 illustrate the normalized hourly demand of the retailer's consumers, the hourly retail prices and the hourly wholesale prices in the examined day, respectively, for different time-shifting flexibility and retailer's relative size scenarios. It should be noted that Fig. 3.13 presents the normalized demand for each scenario regarding α and β ($ND_t^{\alpha,\beta}$) instead of the absolute one, in order to facilitate the subsequent analysis; specifically, the hourly demand in each scenario is normalized with respect to the peak demand in a scenario with the same retailer's relative size β

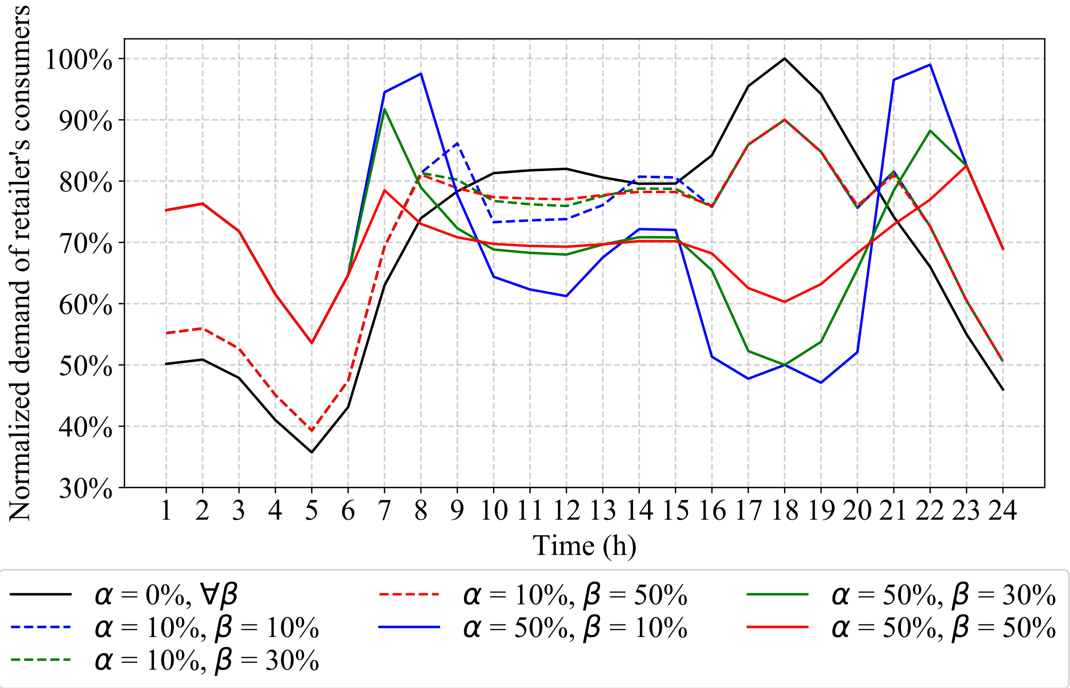

Fig. 3.13 Normalized hourly retail demand for different demand flexibility and retailer's relative size scenarios.

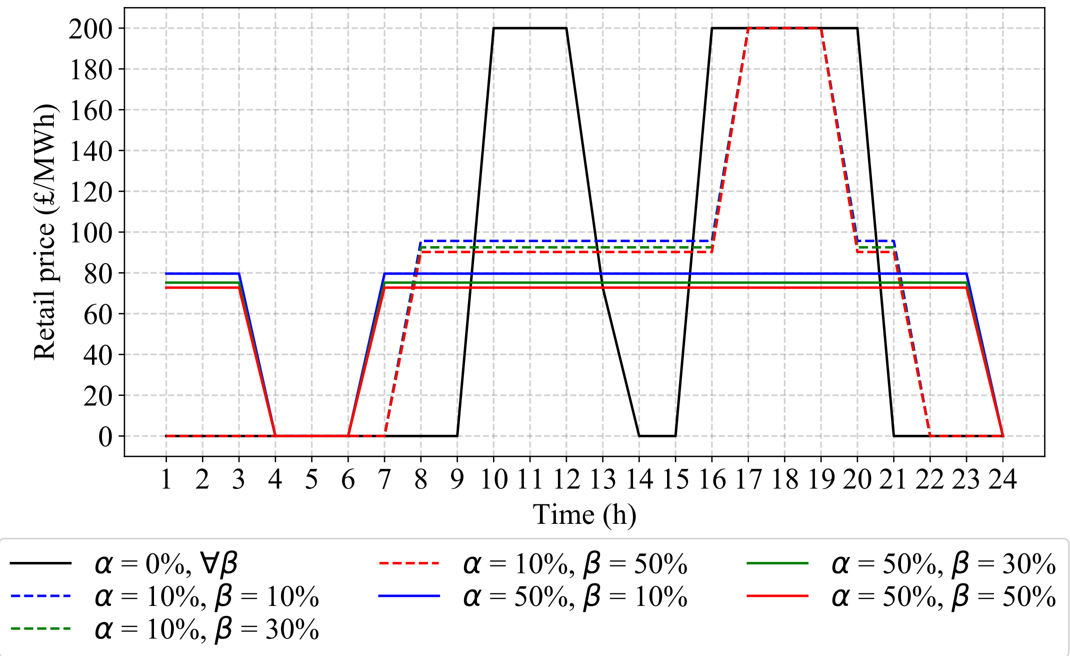

Fig. 3.14 Hourly retail prices for different demand flexibility and retailer's relative size scenarios.

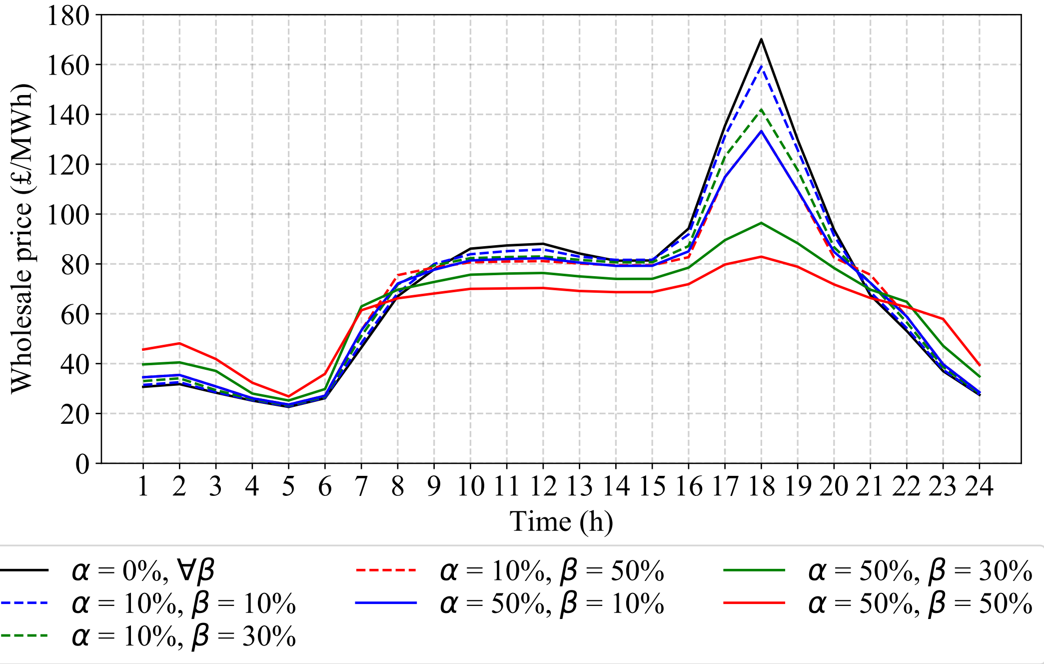

Fig. 3.15 Hourly wholesale market prices for different demand flexibility and retailer's relative size scenarios.

and without demand flexibility ($\alpha = 0\%$):

$$ND_t^{\alpha,\beta} = \frac{demand_t^{\alpha,\beta}}{\max_t demand_t^{\alpha=0\%,\beta}} \tag{3.113}$$

As a result, when the consumers do not exhibit flexibility ($\alpha = 0\%$), the profile of this normalized demand is identical for all relative size scenarios, peaking at 100% (Fig. 3.13). The same trend applies to the retail and wholesale prices (Fig. 3.14 and 3.15), since in the absence of demand flexibility the retailer does not change the offered retail prices and does not affect the wholesale market conditions.

This trend of insensitivity to the retailer's relative size is also observed in the retail prices under a certain (positive) demand flexibility scenario. Specifically, although the shape of the retail prices is affected by the extent of demand flexibility α (a higher α results in a flatter retail price profile, as also explained in Section 3.5.2), it is not substantially affected by the retailer's size β (Fig. 3.14). The only observed impact of β on the retail prices is a small reduction of the retail prices during certain hours as β increases, which is driven by the combination of two factors: i) a higher β enhances the flattening effect of the consumers' demand shifting response on the system demand

and subsequently on the wholesale prices (Fig. 3.15), and also reduces the average wholesale price (Table 3.4), since the impact of the retailer's demand on the wholesale market conditions becomes naturally more prominent when its market share increases, and ii) the average retail price is accordingly reduced (as explained in Section 3.5.2 and demonstrated in Table 3.4). Since this average retail price reduction is relatively small, the observed reduction of the hourly retail prices as β increases is also small.

Table 3.4 Average retail and wholesale price for different demand flexibility and retailer's relative size scenarios.

$\alpha = 0\%$	$\alpha = 10\%$			$\alpha = 50\%$		
$\forall\beta$	$\beta = 10\%$	$\beta = 30\%$	$\beta = 50\%$	$\beta = 10\%$	$\beta = 30\%$	$\beta = 50\%$
69.70	68.85	67.40	66.36	66.36	62.69	60.63

In contrast to the retail prices, the wholesale prices under a certain (positive) demand flexibility scenario are substantially affected by the retailer's size β. Specifically, and as mentioned before, a higher β enhances the flattening effect of the consumers' demand shifting response on the wholesale prices, since the impact of the retailer's demand on wholesale market conditions becomes naturally more prominent when its market share increases. The above discussion implies (rather intuitively) that the retailer's relative size is a factor affecting mostly the interaction between the retailer and the wholesale market and much less the interaction between the retailer and its consumers.

After analyzing the impact of β on retail and wholesale prices, lets us now analyze its impact on the demand of the retailer's consumers; interestingly, a mixed effect is observed in Fig. 3.13, with different trends being evident under different demand flexibility scenarios.

Under a relatively high demand flexibility scenario ($\alpha = 50\%$), the shape of the normalized demand profile is greatly affected by the retailer's size. Specifically, when this size is relatively small ($\beta = 10\%$), the normalized demand profile is somewhat complementary to the benchmark ($\alpha = 0\%$) demand profile and the wholesale price profile, exhibiting increased (with respect to the benchmark profile) demand during hours with low wholesale prices and reduced demand during hours with high wholesale prices. The reason is that, given that the retailer cannot greatly affect the wholesale prices (due to its relatively small β), it prefers to utilize a large extent of the available time-shifting flexibility of its consumers to buy more energy at low-price (off-peak) hours and consequently reduce its total cost. On the other hand, as the retailer's size increases, this complementarity of the normalized demand profile gradually fades and this profile becomes flatter. The reason is that, given that the retailer increasingly affects the wholesale prices (due to its increasing β), it prefers to utilize less the available time-shifting flexibility of its consumers, since otherwise the wholesale prices will become unfavorable, i.e. they will become higher during hours with higher consumers' demand and lower during hours with lower consumers' demand, increasing the retailer's cost in the wholesale market.

In other words, following the discussion in Section 3.5.2, demand flexibility has two (correlated) improving effects on the retailer's cost: a) it enables the retailer to buy more energy at low-price (off-peak) hours and b) reduce the wholesale prices at high-price (peak) hours. This implies that the former effect is more important for the retailer when its size is relatively small, while the latter effect is more important when its size is relatively large.

Moving now to a scenario with a relatively low demand flexibility ($\alpha = 10\%$), the above discussed trends are much less evident and the normalized demand is rather insensitive to the retailer's size (Fig. 3.13). This is because both improving effects of demand flexibility on the retailer's cost are diminished and thus their correlation is not greatly affected by the retailer's size.

The final part of this analysis lies in investigating what are the overall effects of the retailer's relative size on its business case. Fig. 3.16, 3.17 and 3.18 illustrate the % reduction of the retailer's revenue, cost and profit driven by the introduction of relatively low ($\alpha = 10\%$) and relatively high ($\alpha = 50\%$) demand flexibility with respect to the benchmark scenario, for different retailer's relative size scenarios. In other words, the relative revenue reduction ($RRR^{\alpha,\beta}$), relative cost reduction ($RCR^{\alpha,\beta}$) and relative profit reduction ($RPR^{\alpha,\beta}$) in these figures are calculated as:

$$RRR^{\alpha,\beta} = \frac{Revenue_t^{\alpha=0\%,\beta} - Revenue_t^{\alpha,\beta}}{Revenue_t^{\alpha=0\%,\beta}} \, (\%) \tag{3.114}$$

$$RCR^{\alpha,\beta} = \frac{Cost_t^{\alpha=0\%,\beta} - Cost_t^{\alpha,\beta}}{Cost_t^{\alpha=0\%,\beta}} \, (\%) \tag{3.115}$$

$$RPR^{\alpha,\beta} = \frac{Profit_t^{\alpha=0\%,\beta} - Profit_t^{\alpha,\beta}}{Profit_t^{\alpha=0\%,\beta}} \, (\%) \tag{3.116}$$

First of all, under a certain (positive) demand flexibility scenario, the $RRR^{\alpha,\beta}$ is always increased with a larger retailer's size, irrespectively of the extent of this demand flexibility (Fig. 3.16). This trend is driven by the following effects: i) the consumers' (normalized) demand and (to a lesser extent) the offered retail prices become flatter (Fig. 3.13 and Fig. 3.14), and the retailer's revenue is given by the summation of the demand-retail price products across all hours, and ii) the average retail price is reduced (Table 3.4). This trend implies that the deteriorating effect of demand flexibility on the revenue of the retailer (explained in Section 3.5.2) is relatively enhanced as the size increases.

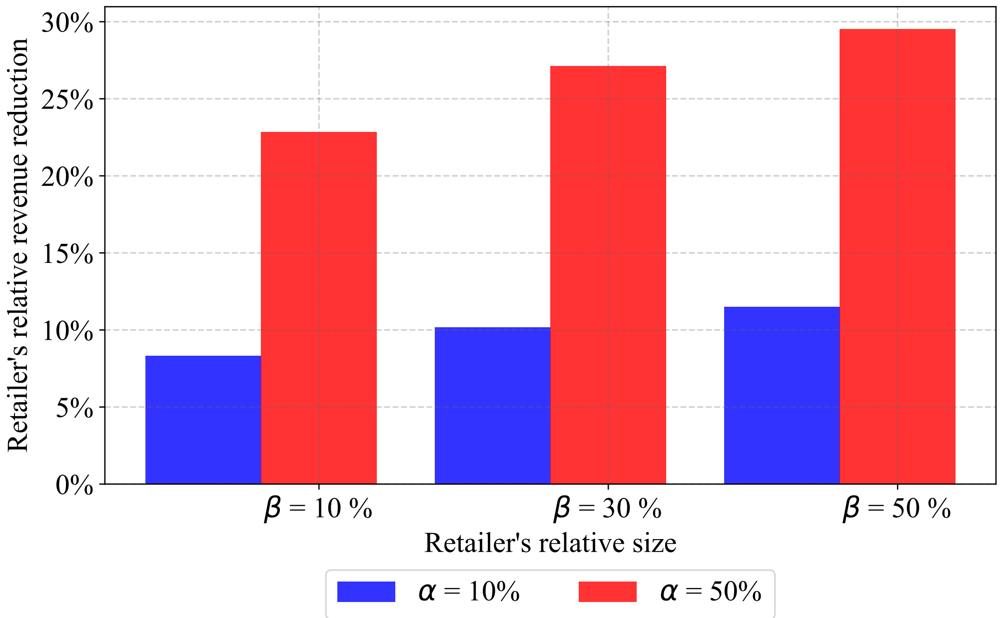

Fig. 3.16 Retailer's relative revenue reduction for different demand flexibility and retailer's relative size scenarios.

Furthermore, under a certain (positive) demand flexibility scenario, the $RCR^{\alpha,\beta}$ is also always increased with a larger retailer's size, irrespectively of the extent of this demand flexibility (Fig. 3.17). This trend is driven by the following effects: i) the consumers' (normalized) demand and the wholesale prices become flatter (Fig. 3.13 and Fig. 3.15), and the retailer's cost is given by the summation of the demand-wholesale price products across all hours, and ii) the average wholesale price is reduced (Table 3.4). This trend implies that the improving effects of demand flexibility on the cost of the retailer (explained in Section 3.5.2) are relatively enhanced as the size of the retailer increases.

In similar vein with the retailer's profit analysis in Section 3.5.2, since the impacts of the retailer's size on its revenue and cost counteract each other, i.e. a larger size increases its revenue reduction (deteriorating its business case) but also increases its

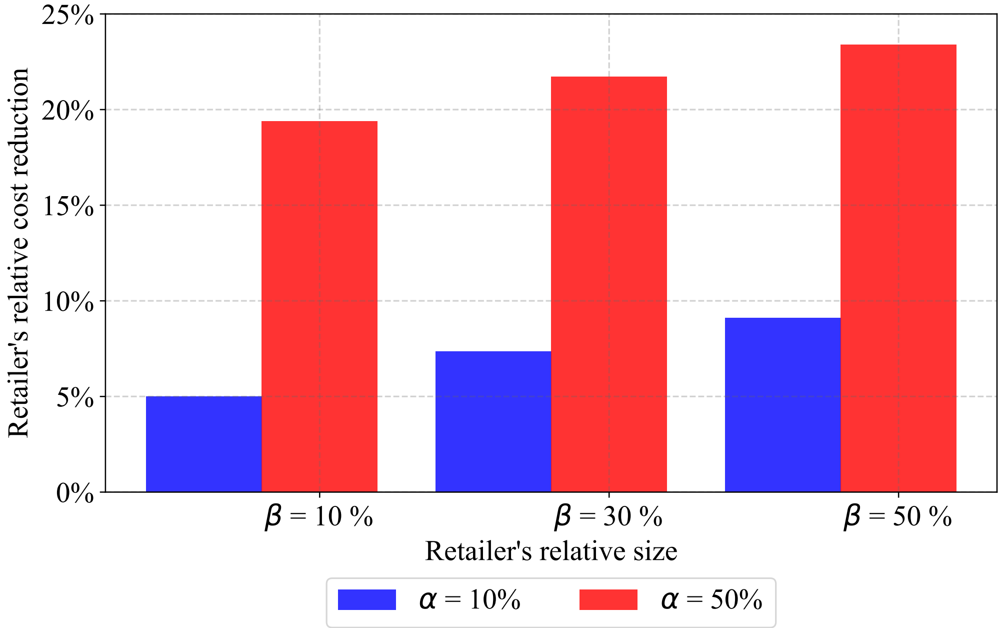

Fig. 3.17 Retailer's relative cost reduction for different demand flexibility and retailer's relative size scenarios.

cost reduction (improving its business case), the overall impact of the retailer's size on its profit depends on which of the above two impacts is more significant. Interestingly, in contrast to the uniform impacts on revenue and cost, a mixed impact is observed in Fig. 3.18, with different trends being evident under different demand flexibility scenarios. This mixed impact of the retailer's size on its profit is linked with the similarly mixed impacts of the retailer's size on the two improving effects of demand flexibility on retailer's cost, discussed earlier in this section.

Under a relatively low demand flexibility scenario ($\alpha = 10\%$), the $RPR^{\alpha,\beta}$ is reduced with a larger retailer's size, which means that a larger size improves the retailer's business case. The reason is that the limited demand flexibility of its consumers does not allow the retailer to significantly reduce its cost in the wholesale market by buying energy at low-price (off-peak) hours. Therefore, in this scenario, the retailer would prefer to have a larger size in order to be able to flatten the wholesale price profile and at least reduce the wholesale prices at high-price (peak) hours (which the retailer cannot greatly avoid due to the limited demand flexibility of its consumers).

Under a relatively high demand flexibility scenario ($\alpha = 50\%$), the $RPR^{\alpha,\beta}$ is increased with a larger retailer's size, which means that a larger size deteriorates the retailer's business case. The reason is that the significant demand flexibility allows the

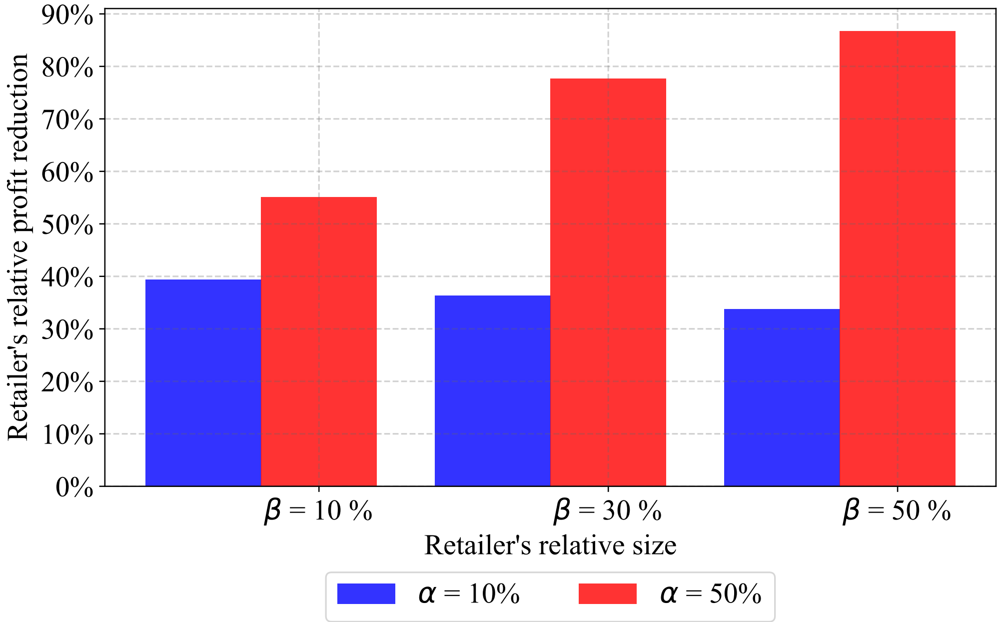

Fig. 3.18 Retailer's relative profit reduction for different demand flexibility and retailer's relative size scenarios.

retailer to significantly reduce its cost in the wholesale market by buying a significant part of its energy requirements at low-price (off-peak) hours. Therefore, in this scenario, the retailer would prefer to have a smaller size in order to avoid flattening the wholesale price profile and increasing the wholesale prices at off-peak hours (which are particularly important due to the significant demand flexibility of its consumers).

3.5.5 Comparison with the State-of-the-art Model

As discussed in Sections 3.2 and 3.3.2, the proposed bi-level optimization model makes a fundamental contribution with respect to the state-of-the-art bi-level models adopted in the existing literature. Specifically, state-of-the-art bi-level models treat the wholesale market prices as exogenous, fixed parameters that are not affected by the retailer's pricing decisions. On the other hand, the proposed bi-level model captures the realistic, indirect impacts of the retailer's pricing decisions on the wholesale prices by representing endogenously the wholesale market clearing process (through the LL2 problem in Fig. 4.1), which becomes particularly important as the size of the retailer increases.

The analysis presented in Sections 3.5.3 and 3.5.4 has demonstrated that this feature of the proposed bi-level model reveals certain important trends around the impact of demand flexibility on the retailer's business, which, on a theoretical basis, cannot be captured through state-of-the-art bi-level models. The aim of this section lies in comparing the state-of-the-art and the proposed modeling approach on a quantitative basis, to clearly demonstrate their differences.

In this context, the state-of-the-art bi-level model has been reproduced by removing the LL2 optimization problem from the mathematical formulation presented in Section 3.4, and assuming that the wholesale prices are exogenous input parameters, which are derived ex-ante by solving the wholesale market clearing problem with fixed (inflexible) demand. Both state-of-the-art and proposed models have been then executed for different scenarios regarding the relative size β of the examined retailer, while assuming that the time-shifting flexibility of the retailer's consumers is $\alpha = 30\%$ and the maximum retail price is $\lambda = 200$£/MWh.

Fig. 3.19, 3.20 and 3.21 illustrate the hourly wholesale prices, the hourly retail prices and the normalized hourly demand of the retailer's consumers in the examined day, for both the state-of-the-art model and the proposed model, and different retailer's relative size scenarios.

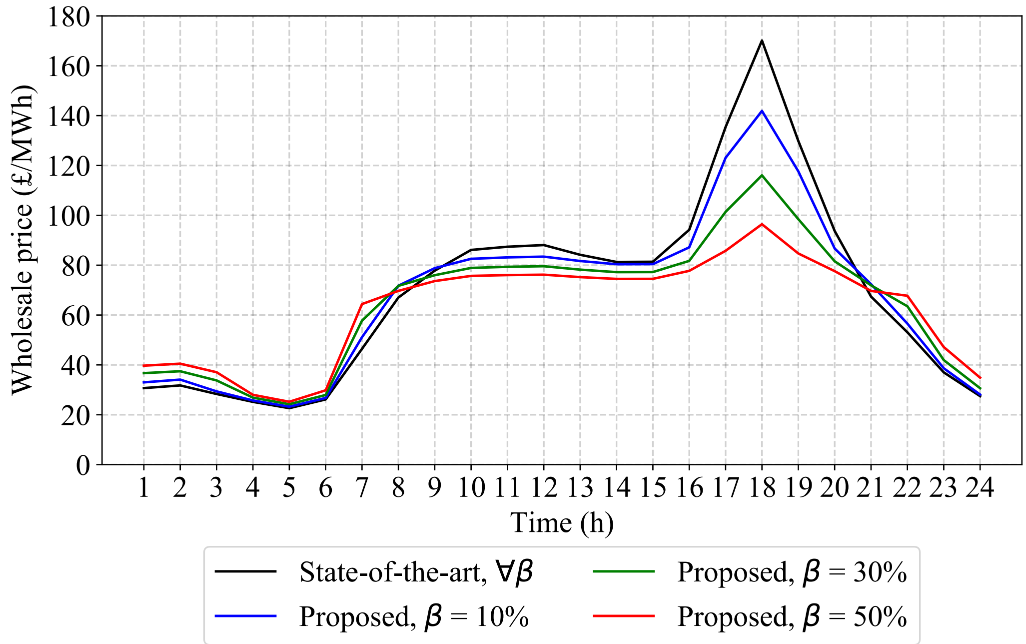

Fig. 3.19 Hourly wholesale market prices for different retailer‘s relative size scenarios under different models.

Given the assumptions of the state-of-the-art model, the wholesale price profile is identical to the respective profile under the benchmark scenario ($\alpha = 0\%$) and

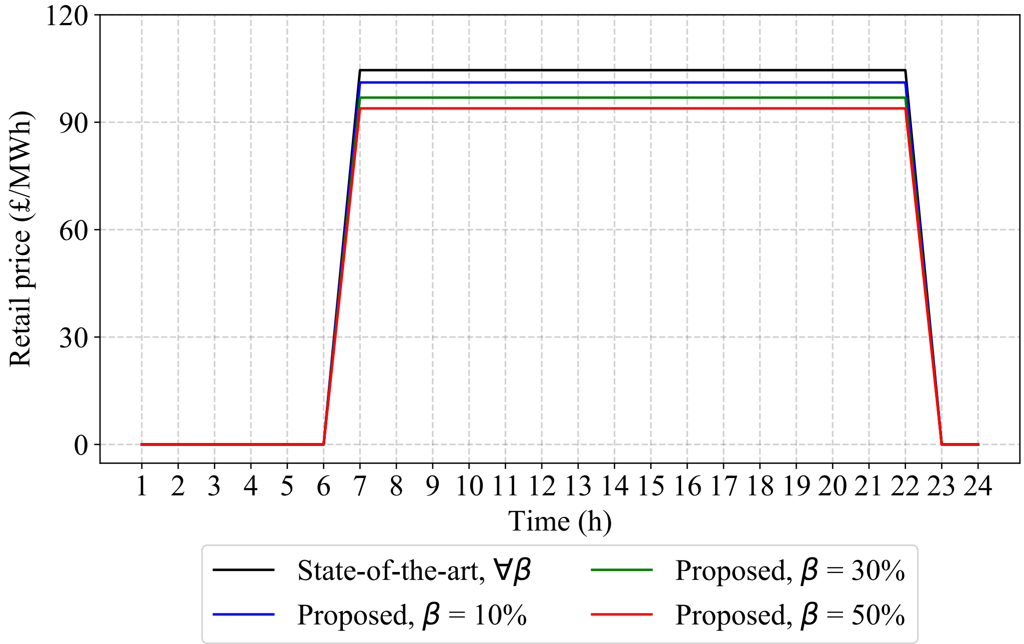

Fig. 3.20 Hourly retail prices for different retailer's relative size scenarios under different models.

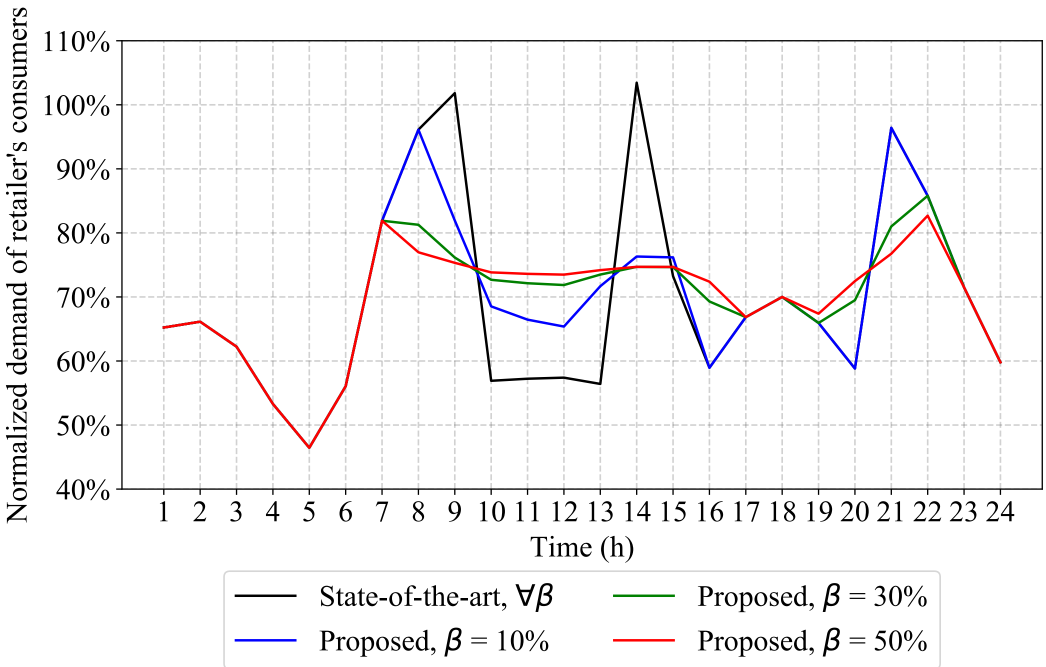

Fig. 3.21 Normalized hourly retail demand for different retailer's relative size scenarios under different models.

is not affected by the retailer's size. Under the proposed model on the other hand, as discussed in Section 3.5.4, the wholesale prices are substantially affected by the retailer's size, with a higher size yielding a flatter wholesale price profile and a lower average wholesale price.

A similar trend applies to the retail prices. Under the proposed model, as discussed in Section 3.5.4, a higher retailer's size reduces (to a small but visible extent) the retail prices during hours outside the off-peak window, driven by the combination of the reduction of the average wholesale price and the regulatory constraint imposed on the average retail price. Under the state-of-the-art model on the other hand, the retail price profile is not affected by the retailer's size, driven by the same insensitivity of the wholesale price profile.

Finally, a similar trend applies to the retail demand. Under the proposed model, as discussed in Section 3.5.4, the shape of the normalized demand profile is greatly affected by the retailer's size. When its size is relatively small ($\beta = 10\%$) and thus cannot greatly affect the wholesale prices, the retailer prefers to utilize a large extent of the available shifting flexibility of its consumers to buy energy at low-price (off-peak) hours, leading to a normalized demand profile which is somewhat complementary to the wholesale price profile. When its size is relatively large ($\beta = 30\%$ and $\beta = 50\%$) and thus substantially affects the wholesale prices, the retailer prefers to utilize less the available shifting flexibility of its consumers (since otherwise the wholesale prices will become unfavorable), leading to a flatter normalized demand profile.

Under the state-of-the-art model on the other hand, it can be observed that: a) the normalized demand profile is not affected by the retailer's size, and b) this profile is complementary to the wholesale price profile, and this complementarity is more acute than the one observed under the proposed model and a relatively small ($\beta = 10\%$) retailer's size, with very high spikes during hours with low wholesale prices. The reason behind these trends is that the retailer does not affect at all the wholesale prices, irrespectively of its size (given the assumptions of the state-of-the-art model), and therefore always prefers to utilize as much as possible the available shifting flexibility of its consumers to buy energy at low-price hours. This discussion implies (rather intuitively) that the state-of-the-art model becomes equivalent to the proposed model and suitable for driving the retailer's decision making, only under the condition that the retailer's size is extremely small.

In order to further demonstrate the differences between the two models, Fig. 3.22 and 3.23 illustrate the retailer's relative revenue reduction $RRR^{\alpha,\beta}$ (defined by (3.114)) and relative cost reduction $RCR^{\alpha,\beta}$ (defined by (3.115)) driven by the introduction of demand flexibility $\alpha = 30\%$ with respect to the benchmark scenario ($\alpha = 0\%$), for both the state-of-the-art model and the proposed model, and different retailer's relative size scenarios.

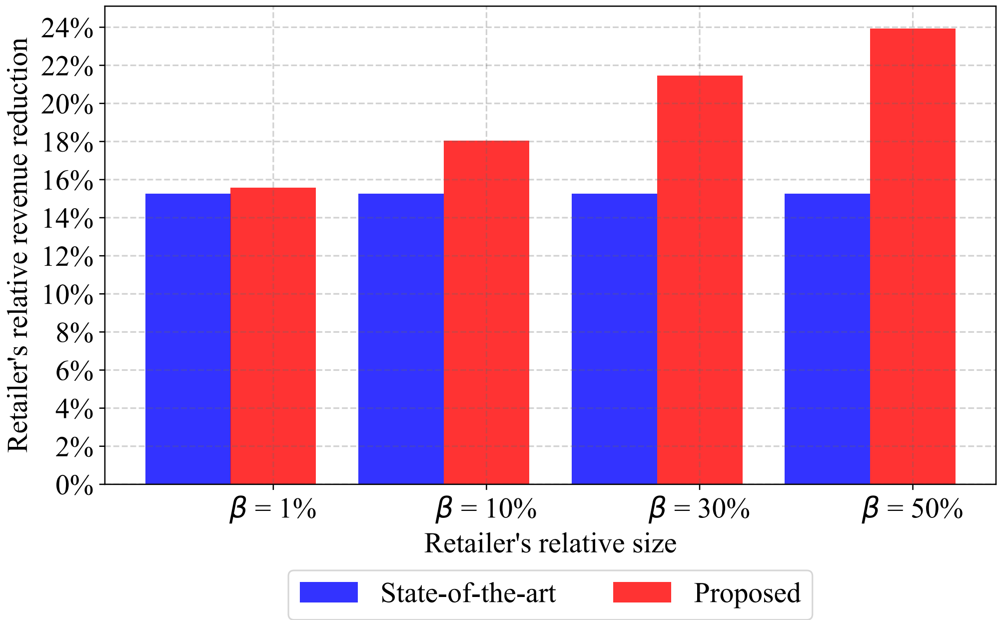

Fig. 3.22 Retailer‘s relative revenue reduction for different retailer's relative size scenarios under different models.

First of all, under the state-of-the-art model, both the $RRR^{\alpha,\beta}$ and the $RCR^{\alpha,\beta}$ are constant and not affected by the retailer's size, which is driven by the same insensitivity of the wholesale prices, retail prices and consumers' demand, as discussed above. Under the proposed model on the other hand, as discussed in Section 3.5.4, both the $RRR^{\alpha,\beta}$ and the $RCR^{\alpha,\beta}$ are increased with a larger retailer's size. This trend implies that the proposed model (in contrast with the state-of-the-art model) captures the impacts of the retailer's size on its business case.

Furthermore, it can be observed that both the $RRR^{\alpha,\beta}$ and the $RCR^{\alpha,\beta}$ are higher under the proposed model than under the state-of-the-art model, for all retailer's size scenarios. The reason behind the higher $RCR^{\alpha,\beta}$ is that, in contrast with the state-of-the-art model, the proposed model captures the impacts of demand flexibility on wholesale prices (flattened profile and reduced average, Fig. 3.19) and therefore captures an additional improving effect of demand flexibility on the retailer's cost (apart from the ability of the retailer to buy energy at more favorable hours, which is captured by both models). The reason behind the higher $RRR^{\alpha,\beta}$ is that the reduction

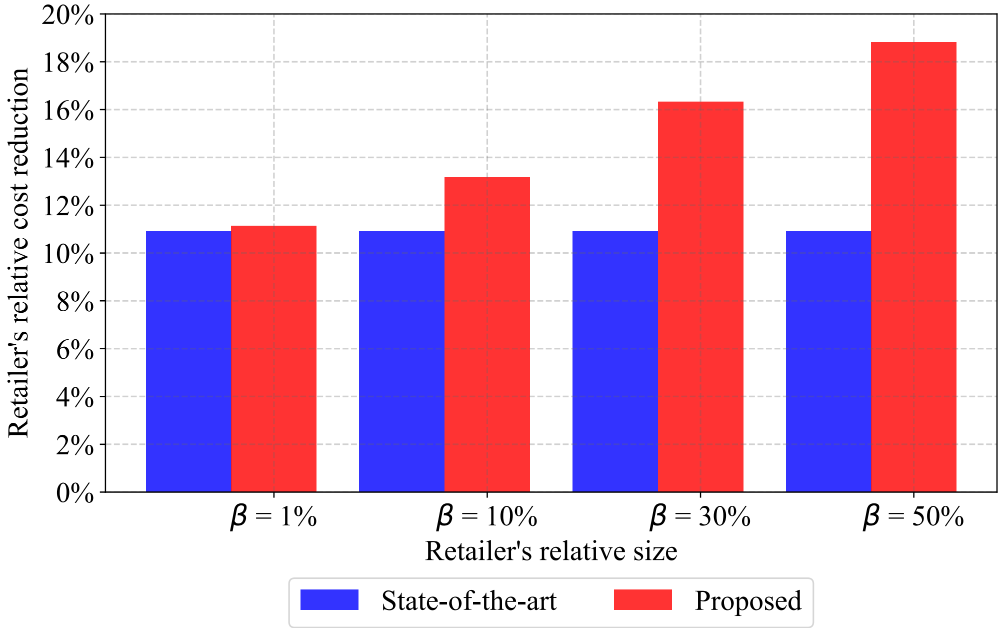

Fig. 3.23 Retailer's relative cost reduction for different retailer's relative size scenarios under different models.

of retail prices (during hours outside the off-peak window) is more significant under the proposed model (Fig. 3.20) due to the reduction of the average wholesale price (which is not captured by the state-of-the-art model) in combination with the regulatory constraint imposed on the average retail price. These trends imply that the state-of-the-art model underestimates both the deteriorating effect of demand flexibility on the retailer's revenue and the improving effect of demand flexibility on the retailer's cost, and thus is not generally suitable for driving the retailer's decision making

Finally, these differences in both the $RRR^{\alpha,\beta}$ and the $RCR^{\alpha,\beta}$ between the two models are enhanced as the retailer's size increases. As previously discussed, this is because the ability of the proposed model to more accurately capture the impacts of demand flexibility on wholesale prices and retail prices becomes more important as the size of the retailer increases. On the other hand, the impact of this feature of the proposed model on the $RRR^{\alpha,\beta}$ and the $RCR^{\alpha,\beta}$ becomes negligible when the size of the retailer becomes very small, as indicated by the scenario $\beta = 1\%$ in Fig. 18 and 19. This validates our previous conclusion that the state-of-the-art model becomes equivalent to the proposed model and suitable for driving the retailer's decision making, only under the limiting condition that the retailer's size is extremely small.

3.6 Conclusions

This chapter has proposed a novel bi-level optimization model to comprehensively investigate the effects of demand flexibility on an electricity retailer's business case. In contrast with state-of-the-art bi-level optimization models, this model drops the unrealistic assumption that the retailer treats wholesale market prices as exogenous, fixed parameters, and represents endogenously the wholesale market clearing process, thus capturing the realistic implications of the retailer's pricing strategies on the wholesale market prices.

The presented case studies provide numerous new and valuable insights. First of all, they demonstrate the interactions between the retailer, the flexible consumers and the wholesale market and demonstrate that demand flexibility: a) reduces the retailer's revenue from the consumers, since it limits the retailer's strategic potential of exploiting the consumers through setting large retail price differentials between peak and off-peak hours, b) also reduces the retailer's cost in the wholesale market, since it enables the retailer to buy more energy at low-price hours and also reduce the wholesale prices at high-price hours, and c) does not have a uniform impact on the retailer's overall profit.

Going further, this impact of demand flexibility on the retailer's profit is shown to depend on the strictness of the regulatory constraints imposed on the offered retail prices. Under a looser regulatory framework demand flexibility reduces the overall profit of the retailer, while under a stricter regulatory framework it increases this profit.

Moreover, the presented case studies analyse how the above effects of demand flexibility depend on the retailer's relative size in the market. Specifically, the obtained results demonstrate that both the deteriorating effect of demand flexibility on the retailer's revenue and its improving effect on the retailer's cost are relatively enhanced as the size of the retailer increases. However, the implications of the retailer's size on its overall profit are not uniform, but depend on the extent of demand flexibility. Under relatively high demand flexibility the retailer achieves a higher profit when its size is smaller, while under relatively low demand flexibility the retailer achieves a higher profit when its size is larger.

Finally, the presented case studies highlight the added value of the proposed bi-level model by comparing its outcomes against the state-of-the-art bi-level model. Specifically, the obtained results demonstrate that the state-of-the-art bi-level model underestimates the above effects of demand flexibility on the retailer's revenue and cost, compared to the proposed model, and this underestimation is enhanced as the retailer's size increases. This result implies that the state-of-the-art model is suitable for driving a retailer's decision-making only under the limiting condition that the retailer's size is extremely small (around 1% of the market according to the obtained results), in contrast with the general suitability of the proposed model.

Given the conclusion of physical insights from the presented case studies, the policy recommendations are also highlighted to promote this novel work. First of all, demand flexibility is highly encouraged to reduce the energy cost both in retail and wholesale markets. Secondly, demand flexibility can effectively complement regulatory policies in safeguarding the consumers against the strategic behavior of retailer. Finally, the small and new players in the retail market are more likely to take initiatives towards the realization of the flexibility potential of their consumers, than large and incumbent retailers.

Chapter 4

Strategic Pricing for Local Energy Market

4.1 Introduction

Driven by environmental concerns, governments worldwide have taken significant initiatives towards the decarbonization of energy systems, mainly involving the large-scale integration of renewable generation and the electrification of transport and heat sectors of the demand side [1]. However, these initiatives introduce critical techno-economic challenges to electricity systems' operation and planning; conventional generators incur significant costs to provide system balancing services given the inherent variability of renewable generation, while electric transport and heat loads drive capital-intensive generation and network reinforcements [94].

A very promising paradigm change towards addressing these challenges and achieving a cost-effective transition to the low-carbon future lies in the deployment of flexible, small-scale distributed energy resources (DER), including flexible loads, micro-generators and energy storage. These DER exhibit significant operating flexibility which can support system balancing and reduce demand peaks, thereby limiting the balancing costs of conventional generators and the investments costs of new generation and network assets. In other words, this paradigm change partially shifts the provision of security of supply and system services from large-scale centralised assets to small-scale DER. Numerous recent works have demonstrated and analyzed the value of this paradigm change [94, 95, 5, 96]. However, this paradigm change greatly complicates the operation of the system, as the effective coordination of large numbers of small-scale DER involves very significant communication and computational scalability challenges as well as privacy concerns, since DER owners are not generally willing to disclose private information and be directly controlled by external entities [97].

One of the coordination approaches that has recently attracted great interest by both industry and academia is the establishment of local energy markets (LEM). Under this approach, the global coordination burden is broken down to the coordination of local market clusters, each grouping a number of end users with DER (lately referred to as community), and enabling direct energy trading between them [98]. Beyond addressing scalability and privacy challenges, the LEM concept exhibits a number of potential benefits, such as a) deferring or avoiding distribution network reinforcements as a result of matching local demand with local generation, b) enhancing the engagement of local end-users in system operation by creating a local identity and promoting social cooperation, and c) revitalizing the local economy by shaping opportunities for local investment, creating new jobs at the community level and promoting self-sufficiency. Recent studies have illustrated these benefits of LEM [99–101].

However, the introduction of LEM is also expected to have significant implications on the strategic interactions between the local end users and incumbent electricity retailers. After the deregulation of electricity systems, electricity retailers have been representing the large majority of the end users in the wholesale electricity market, by buying / selling energy consumed / produced by their end users from / to the wholesale market. The objective of these self-interested market entities lies in maximizing their individual profits by optimizing the retail prices offered to their customers but also respecting regulatory limits [8]. With the introduction of LEM, customers can directly trade energy between them, thus limiting their energy dependency on the incumbent retailer and increasing their economic surplus. This effect is expected to in turn impact the retailer's decisions regarding the offered prices and eventually its profit.

4.2 Literature Review

Numerous previous works in the existing literature have investigated the interactions between strategic retailers and individual customers with different types of DER [22–24, 27, 28, 102, 30, 36, 37, 33, 103]. Table 4.1 summarizes in a structured way the main characteristics of these works. The representation of customers' flexibility in these papers includes flexible loads, micro-generation, energy storage. Furthermore, these papers generally adopt two different methodologies to model the interaction between an electricity retailer and its flexible customers.

Table 4.1 Summary of existing literature associated with the examined problem.

Paper	Optimization model	Approach	DER type	Local energy market
[22]	Single-level	Elasticity function	Flexible demand	No
[23]	Single-level	Elasticity function	Flexible demand	No
[24]	Single-level	Elasticity function	Flexible demand	No
[27]	Single-level	Elasticity function	Flexible demand, Wind Distributed generation, PV Energy storage system	No
[28]	Single-level	Elasticity function	Flexible demand, Wind Distributed generation, PV Energy storage system, EV	No
[102]	Bi-level	MPEC	Distributed generation	No
[30]	Bi-level	MPEC	Flexible demand	No
[36]	Bi-level	Iterative algorithm	Flexible demand	No
[37]	Bi-level	Iterative algorithm	Flexible demand	No
[33]	Bi-level	MPEC	Flexible demand Distributed generation	No
[103]	Bi-level	MPEC	Flexible demand Distributed generation	No
This work	Bi-level	MPEC Binary relaxation Primal-dual	Flexible demand Distributed generation Energy storage system	Yes

Specifically, papers [22–24, 27, 28] employ single-level optimization models, which aim at maximizing the retailer's profit; in these models, the optimal response of the flexible DERs to the retail prices is expressed as closed-form elasticity functions. The drawback of this approach lies in the fact that the parameters of these functions are determined based on exogenous data and therefore cannot accurately capture the impacts of the retail prices on the DERs' response.

In order to address this limitation, papers [102, 30, 36, 37, 33, 103] employ bi-level optimization models in order to rigorously capture the interactions between the optimization of the pricing decisions of the self-interested retailer (modeled in the upper level – UL) and the optimization of the demand response of its flexible DERs (modeled in the lower level – LL), which is thus represented endogenously and not based on exogenous data. These bi-level optimization problems are solved either through converting them into Mathematical Programs with Equilibrium Constraints (MPEC) or by employing iterative approaches (Table 4.1).

Although the existing literature has investigated the strategic interactions between retailers and individual customers with different types of DER [22–24, 27, 28, 102, 30, 36, 37, 33, 103]. Table 4.1, the respective interactions between retailers and LEM integrating all these DER types and enabling energy trading between them, have not yet been analytically explored.

4.3 Approach

This chapter makes the first attempt to fill this knowledge gap by proposing a novel multi-period bi-level optimization problem that captures the interactions between the retailer's strategic pricing decisions formulated in the UL problem and the operation of an LEM among its served customers formulated in the LL problem.

Since the LL problem is non-convex, as it includes the binary decision variables of the LEM to either buy or sell energy to the retailer at each period, the formulated bi-level optimization problem cannot be solved through the traditional approach of converting it to a Mathematical Program with Equilibrium Constraints (MPEC). In this context, we develop a new approach, which is based on the relaxation and primal-dual reformulation of the original, non-convex LL problem and the penalization of the associated duality gap.

4.3.1 Modeling Assumptions

For clarity reasons, the main assumptions behind the proposed model are outlined below:

1. The customers served by the examined retailer are modelled with different types of DER (flexible loads, micro-generation, energy storage), and each customer receiving the retail prices signals operates under its individual operating characteristics.

2. The decision making problem of the examined retailer considers both the interaction with the served independent customers (i.e. the flexible consumers and energy storages to which it sells energy, and the micro-generators and energy storages from

which it buys energy) and the interaction with the LEM (to which it sells energy or from which it buys energy), with the overall aim of maximizing the retailer's profit.

3. The examined retailer is assumed to serve a relatively small population of customers and therefore its decisions do not affect the wholesale prices, which are thus treated as fixed, exogenous parameters of the problem, i.e. the retailer is assumed to be a price-taker in the wholesale market.

4. In contrast with the traditional fixed pricing or time-of-use pricing regimes where the offered retail prices are flat throughout the examined daily horizon or during certain intervals of this horizon (e.g. peak and off-peak periods), the examined strategic retailer can offer hour-specific retail prices to the served customers. In order to prevent the retailer from exploiting the customers and making excessive profits, regulatory constraints are imposed on the maximum retail prices it can offer to its customers [103].

5. Reflecting the reality, it is assumed that the examined retailer can offer differentiated prices to its customers for buying and selling energy, but these buy and sell prices are not differentiated for different customers. Such a kind of dual pricing mechanism means that the strategic retailer is capable of offering a relative high buy price and a relative low sell price in order to obtain more revenue from but pay less cost to its served local customers. This assumption also allows its served customers to have incentive to participate in the local energy market when they see the price differentials between retail buying and retail selling.

6. For presentation clarity and without loss of generality, it is also assumed that only one LEM operates among (a subset of) the retailer's served customers; each customer is assumed to either participate in the LEM or trade independently with the retailer, neglecting the potential of customers participating directly in the wholesale market (which is a realistic assumption for small size customers). It would be also interesting to model how the customers are optimized to choose between retailer and LEM, which is also one extension of the future work.

4.3.2 Structure of the Bi-level Optimization Model

In order to comprehensively capture the interactions between the retailer, its individual customer, and the LEM, the proposed model is formulated as a bi-level optimization problem, the structure is illustrated in Fig. 4.1.

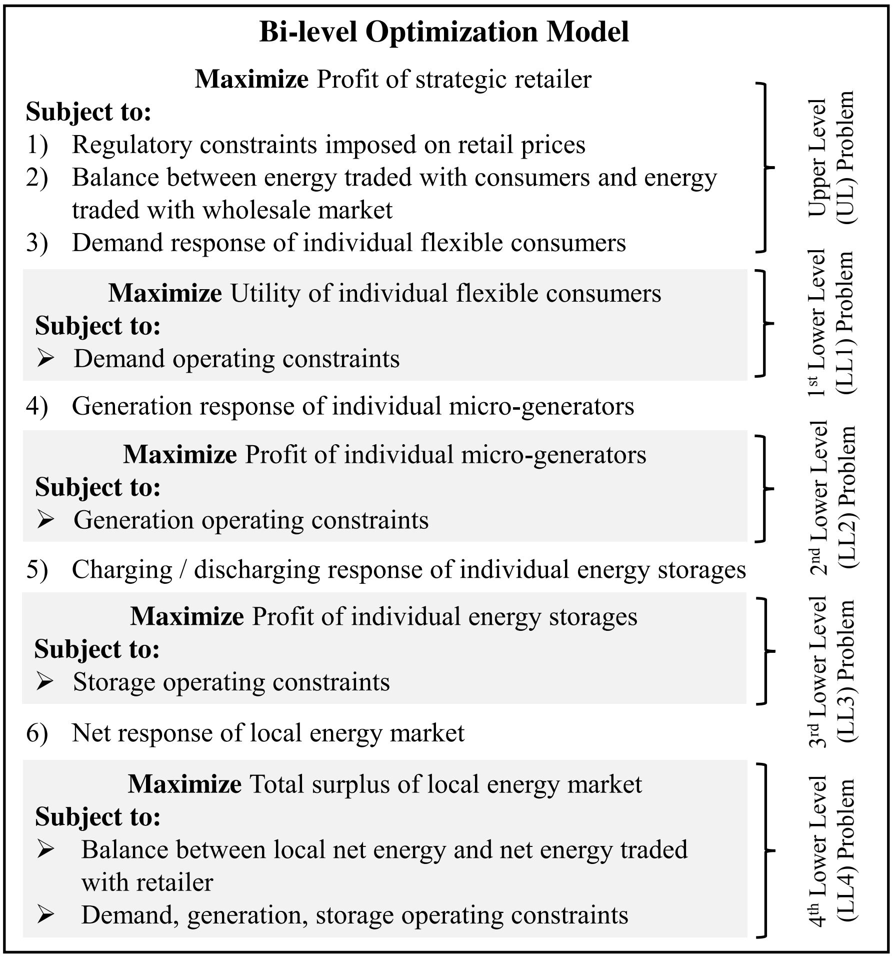

Fig. 4.1 Structure of proposed bi-level optimization model.

1. This bi-level model consists of an upper-level (UL) problem and four lower-level (LL) problems.

2. The UL problem represents the strategic decision making of a self-interested retailer who determines the optimal time-specific retail prices offered to the customers for buying and selling energy so as to maximize its profit, while respecting i) the regulatory constraints imposed on the offered retail prices, and ii) the balance constraint between the energy sold / purchased to / from its served customer as well as the LEM and the energy bought / sold from / to the wholesale market at each time period.

3. This UL problem is subject to four LL problems. As shown in Fig. 1, those four LL problems respectively represent:

(a) The first three of them (LL1-3) represent the decision making of flexible consumers (FC), micro-generators (MG) and energy storages (ES), which do not participate in the LEM and thus optimize their demand / generation response to the offered retail prices so as to maximize their own economic surplus, and subject to their individual operation constraints;

(b) The fourth LL (LL4) problem represents the operation of an LEM with different participants (FC, MG and ES) which, given the offered retail prices and the techno-economic parameters of its participants, determines the optimal dispatch of these participants and the energy exchanges with the retailer so as to maximize their total surplus, and subject to their individual operation constraints.

4. The upper-level problem and the two lower-level problems of Fig. 4.1 are all coupled as illustrated in Fig. 4.2, since:

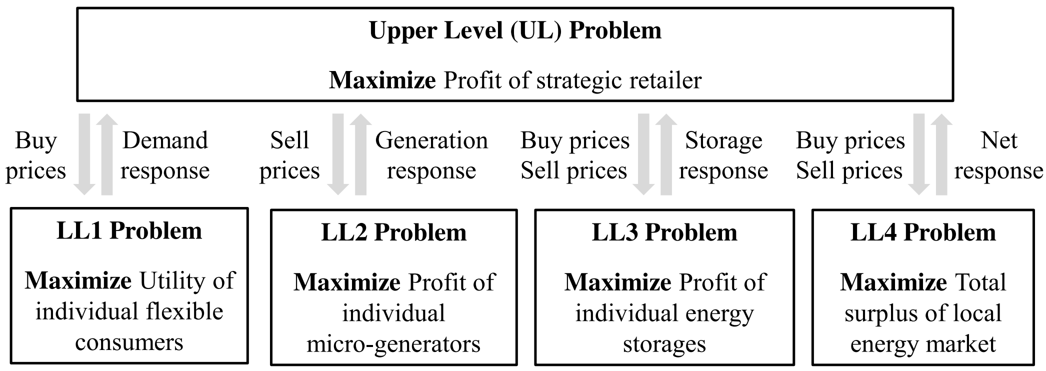

Fig. 4.2 Interrelation between the upper-level and 4 lower-level problems.

(a) The buy and sell prices determined by the retailer (UL) affect the responses of the customers (as they constitute part of the objective functions of the LL1-LL4 problems);

(b) These responses affect the retailer's pricing decisions (as they constitute part of both the objective function and the energy balance constraints of the UL problem).

4.4 Formulation

4.4.1 Decision Making of Strategic Retailer

The upper-level (UL) problem optimizes the pricing decisions of the examined retailer and is formulated as follows:

$$\max_{V^{UL}=\{\lambda_t^b,\lambda_t^s,w_t\}} \sum_t \lambda_t^b \Big(\sum_i d_{i,t} + \sum_k s_{k,t}^c + u_t n_t \Big) - \\ \sum_t \lambda_t^s \Big(\sum_j g_{j,t} + \sum_k s_{k,t}^d + (u_t - 1) n_t \Big) - \sum_t \lambda_t^w w_t \tag{4.1}$$

subject to:

$$\lambda^{min} \leq \lambda_t^b \leq \lambda^{max}, \forall t \tag{4.2}$$

$$\lambda^{min} \leq \lambda_t^s \leq \lambda^{max}, \forall t \tag{4.3}$$

$$\sum_i d_{i,t} - \sum_j g_{j,t} + \sum_k (s_{k,t}^c - s_{k,t}^d) + n_t = w_t, \forall t \tag{4.4}$$

The objective function (4.1) maximizes the overall profit of the retailer, which includes the following components: i) its revenue from selling energy to demanding customers, including independent FC, independent ES when charging, and the LEM when buying energy from the retailer (first term), ii) its cost of buying energy from generating customers, including independent MG, independent ES when discharging, and the LEM when selling energy to the retailer (second term); and iii) its net cost in the wholesale market, i.e. its cost / revenue of buying / selling energy from / to the wholesale market (third term).

The offered retail prices are subject to the regulatory constraints (4.2)-(4.3) which aim at preventing the retailer from exploiting the customers and making excessive profits. Constraints (4.4) express the energy balance constraints of the retailer; the net energy traded with its customers (including independent FC, independent MG, independent ES and the LEM) and the net energy traded with the wholesale market are equal at each period.

4.4.2 Demand Response of Individual Flexible Consumers

The first LL problem (LL1) optimizes the demand response of independent FC (not participating in the LEM) to the retail (buy) prices determined by the retailer, and is formulated as follows for independent FC i:

$$\max_{P^{LL1}=\{d_{i,t}\}} \Big(\sum_t (l_{i,t}^D \, d_{i,t} - q_{i,t}^D \, d_{i,t}^2) - \sum_t \lambda_t^b \, d_{i,t} \Big) \tag{4.5}$$

subject to:

$$0 \leq d_{i,t} \leq d_{i,t}^{max} : \mu_{i,t}^{d-}, \mu_{i,t}^{d+}, \forall t \tag{4.6}$$

The objective function (4.5) maximizes the utility of independent FC i, which is given by the difference between i) its benefit (or satisfaction) perceived from the use of energy (first term), modelled as a quadratic function of its demand levels; and ii) its cost of buying energy from the retailer (second term), modelled as the product of retail (buy) prices and its demand levels. Constraint (4.6) expresses the flexibility of independent FC i to modify its demand within certain limits.

4.4.3 Generation Response of Independent Micro-Generators

The second LL problem (LL2) optimizes the generation response of independent MG (not participating in the LEM) to the retail (sell) prices determined by the retailer, and is formulated as follows for independent MG j:

$$\max_{P^{LL2}=\{g_{j,t}\}} \Big(\sum_t \lambda_t^s \, g_{j,t} - \sum_t (l_j^G \, g_{j,t} + q_j^G \, g_{j,t}^2) \Big) \tag{4.7}$$

subject to:

$$0 \leq g_{j,t} \leq g_j^{max} : \mu_{j,t}^{g-}, \mu_{j,t}^{g+}, \forall t \tag{4.8}$$

The objective function (4.7) maximizes the profit of independent MG j, which is given by the difference between i) its revenue from selling energy to the retailer (first term), modelled as a quadratic function of its power outputs; and ii) its cost of production (second term), modelled as the product of retail (sell) prices and its power outputs. Constraint (4.8) expresses the power output limits of independent MG j.

4.4.4 Charging / Discharging Response of Independent Energy Storages

The third LL problem (LL3) optimizes the charging / discharging response of independent ES (not participating in the LEM) to the retail (buy and sell) prices determined by the retailer, and is formulated as follows for independent ES k:

$$\max_{P^{LL3}=\{s^c_{k,t},\,s^d_{k,t},\,E_{k,t}\}} \Big(\sum_t \lambda^s_t \, s^d_{k,t} - \sum_t \lambda^b_t \, s^c_{k,t} \Big) \tag{4.9}$$

subject to:

$$E_{k,t} = E_{k,t-1} + s^c_{k,t}\,\eta^c_k - s^d_{k,t}/\eta^d_k : \xi_{k,t}, \forall t \tag{4.10}$$

$$E^{min}_k \leq E_{k,t} \leq E^{max}_k : \mu^{se-}_{k,t}, \mu^{se+}_{k,t}, \forall t \tag{4.11}$$

$$0 \leq s^c_{k,t} \leq s^{max}_k : \mu^{sc-}_{k,t}, \mu^{sc+}_{k,t}, \forall t \tag{4.12}$$

$$0 \leq s^d_{k,t} \leq s^{max}_k : \mu^{sd-}_{k,t}, \mu^{sd+}_{k,t}, \forall t \tag{4.13}$$

$$E^0_k = E_{k,NT} : \xi^0_k \tag{4.14}$$

The objective function (4.9) maximizes the profit of independent ES k, which is given by the difference between i) its revenue from selling energy to the retailer when discharging (first term) and ii) the cost of buying energy from the retailer when charging (second term). Constraint (4.10) expresses the energy balance of the ES including charging and discharging losses. Constraint (4.11)-(4.13) represents its minimum and maximum energy and power limits. Finally, constraint (4.14) expresses the energy neutrality assumption, i.e. the ES energy content at the start and the end of the examined horizon are assumed equal.

4.4.5 Operation of Local Energy Market

Beyond individually contracted FC, MG and ES, the examined retailer trades energy with a LEM generally including the same types of customers and enabling energy trading between its participating customers. The fourth LL problem (LL4) represents the operation of the LEM which optimizes the dispatch of the participating customers and the energy exchanges with the retailer, and is formulated as follows:

$$\max_{P^{LL4}} \Big(\sum_{i',t} (l^D_{i',t}\, d_{i',t} - q^D_{i',t}\, d^2_{i',t}) - \sum_{j',t} (l^G_{j'}\, g_{j',t} + q^G_{j'}\, g^2_{j',t}) - \sum_t \lambda^b_t\, u_t\, n_t + \sum_t \lambda^s_t\, (u_t - 1)\, n_t \Big) \tag{4.15}$$

where:

$$P^{LL4} = \{d_{i',t},\, g_{j',t},\, s^c_{k',t},\, s^d_{k',t},\, E_{k',t},\, u_t,\, n_t\} \tag{4.16}$$

subject to:

$$\sum_{i'} d_{i',t} - \sum_{j'} g_{j',t} + \sum_{k'} (s^c_{k',t} - s^d_{k',t}) = n_t : \lambda^l_t, \forall t \tag{4.17}$$

$$u_t \in \{0,1\}, \forall t \tag{4.18}$$

$$0 \leq d_{i',t} \leq d_{i',t}^{max} : \mu_{i',t}^{d-}, \mu_{i',t}^{d+}, \forall i', \forall t \tag{4.19}$$

$$0 \leq g_{j',t} \leq g_{j'}^{max} : \mu_{j',t}^{g-}, \mu_{j',t}^{g+}, \forall j', \forall t \tag{4.20}$$

$$E_{k',t} = E_{k',t-1} + s_{k',t}^{c}\,\eta_{k'}^{c} - s_{k',t}^{d}/\eta_{k'}^{d} : \xi_{k',t}, \forall k', \forall t \tag{4.21}$$

$$E_{k'}^{min} \leq E_{k',t} \leq E_{k'}^{max} : \mu_{k',t}^{se-}, \mu_{k',t}^{se+}, \forall k', \forall t \tag{4.22}$$

$$0 \leq s_{k',t}^{c} \leq s_{k'}^{max} : \mu_{k',t}^{sc-}, \mu_{k',t}^{sc+}, \forall k', \forall t \tag{4.23}$$

$$0 \leq s_{k',t}^{d} \leq s_{k'}^{max} : \mu_{k',t}^{sd-}, \mu_{k',t}^{sd+}, \forall k', \forall t \tag{4.24}$$

$$E_{k'}^{0} = E_{k',NT} : \xi_{k'}^{0}, \forall k' \tag{4.25}$$

The objective function (4.15) maximizes the total surplus of the LEM, which includes the following components: i) the total benefit of all participating FC i' (first term), ii) the total production cost of all participating MG j' (second term), iii) the cost of buying energy from the retailer when the LEM exhibits excess demand (third term), and iv) the revenue from selling energy to the retailer when the LEM exhibits excess generation (fourth term).

Constraints (4.17) express the energy balance constraints of the LEM, ensuring that the excess demand / generation is bought from / sold to the retailer. The fact that the LEM can either buy energy from the retailer or sell energy to the retailer at each period is expressed through the binary decision variable u_t (4.18). When the LEM exhibits excess demand ($n_t > 0$), it needs to buy energy from the retailer ($u_t = 1$), implying that the LEM incurs a cost to the retailer (as the third term of (4.15) becomes positive) and does not earn a revenue from the retailer (as the fourth term of (4.15) becomes zero). On the other hand, when the LEM exhibits excess generation ($n_t < 0$), it needs to sell energy to the retailer ($u_t = 0$), implying that the LEM earns a revenue from the retailer (as the fourth term of (4.15) becomes positive) and does not incur a cost to the retailer (as the third term of (4.15) becomes zero). Finally, constraints (4.19)-(4.25) express the individual operating constraints of the FC, MG and ES participating in the LEM, and follow the formulations adopted in LL1-LL3 problems, respectively.

4.4.6 Reformulation of Non-convex LL problem

The traditional approach for solving such a bi-level optimization problem in the existing literature (see for example [102, 30, 33, 103] lies in converting it to a single-level Mathematical Program with Equilibrium Constraints (MPEC). This is achieved by replacing the LL problems by their equivalent Karush-Kuhn-Tucker (KKT) optimality conditions, provided that the LL problems are continuous and convex, and introducing these conditions as constraints of the UL problem.

However, in the examined bi-level optimization problem, the LL4 problem representing the operation of the LEM, is non-convex as it includes the binary decision variables u_t of the LEM to either buy energy from the retailer or sell energy to the retailer at each period (4.18). This non-convexity prevents the derivation of equivalent KKT conditions for LL4 and thus means that the traditional MPEC approach is not applicable to the examined problem to guarantee its optimality.

In order to address this fundamental challenge and solve the examined problem, this chapter employs a new approach recently proposed by the authors in [104].

Relax Non-convex LL4 problem

Problem LL4 is converted from a non-convex mixed-integer problem to a convex problem by relaxing its binary constraints (4.18) as continuous constraints:

$$0 \leq u_t \leq 1 : \mu_t^{u-}, \mu_t^{u+}, \forall t \tag{4.26}$$

with μ_t^{u-}, μ_t^{u+} being their respective dual variables. This relaxation enables the definition of dual variables for all the constraints of the LL4 problem, which are indicated after a colon in constraints (4.17) and (4.19)-(4.25) above. In other words, this relaxation coverts the LL4 problem from a mixed-integer linear problem to a continuous linear problem, defined by:

$$\max_{V^P} \text{(LL4 objective function)} : (4.15) \tag{4.27}$$

subject to:

$$\text{(LL4 continuous constraints)} : (4.17) \text{ and } (4.19) - (4.25) \tag{4.28}$$

$$\text{(LL4 relaxed binary constraint)} : (4.26) \tag{4.29}$$

Dual Formulation of Relaxed LL4 problem

After the above conversion, the dual problem associated with the relaxed LL4 problem (4.27)-(4.29) can be derived, which is formulated as follows:

$$\max_{D^{LL4}} \Big(-\sum_{i',t} \mu^{d+}_{i',t} d^{max}_{i',t} - \sum_{j',t} \mu^{g+}_{j',t}\, g^{max}_{j'} + \sum_{k',t} \mu^{se-}_{k',t}\, E^{min}_{k'} - \sum_{k',t} \mu^{se+}_{k',t}\, E^{max}_{k'} - \sum_{k',t} \mu^{sc+}_{k',t}\, s^{max}_{k'} - \sum_{k',t} \mu^{sd+}_{k',t}\, s^{max}_{k'} - \sum_{k',(t=1)} \xi_{k',t}\, E^{0}_{k'} - \sum_{k'} \xi^{0}_{k'}\, E^{0}_{k'} - \sum_{t} \mu^{u+}_{t} \Big) \tag{4.30}$$

where:

$$D^{LL4} = \{\lambda^l_t,\, \mu^{d+}_{i',t},\, \mu^{g+}_{j',t},\, \xi_{k',t},\, \mu^{se-}_{k',t},\, \mu^{se+}_{k',t},\, \mu^{sc+}_{k',t},\, \mu^{sd+}_{k',t},\, \xi^{0}_{k'},\, \mu^{u+}_{t}\} \tag{4.31}$$

subject to:

$$\frac{\partial \mathcal{L}_{LL4}}{\partial d_{i',t}} = \lambda^l_t - (l^D_{i',t} - 2\, q^D_{i',t}\, d_{i',t}) + \mu^{d+}_{i',t} \geq 0, \forall i', \forall t \tag{4.32}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial g_{j',t}} = (l^G_{j'} - 2\, q^G_{j'} g_{j',t}) - \lambda^l_t + \mu^{g+}_{j',t} \geq 0, \forall j', \forall t \tag{4.33}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial s^c_{k',t}} = \lambda^l_t - \zeta_{k',t}\, \xi^c_{k'} + \mu^{sc+}_{k',t} \geq 0, \forall k', \forall t \tag{4.34}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial s^d_{k',t}} = -\lambda^l_t + \xi_{k',t}/\eta^d_{k'} + \mu^{sd+}_{k',t} \geq 0, \forall k', \forall t \tag{4.35}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial E_{k',t}} = \zeta_{k',t} - \zeta_{k',t+1} - \mu^{se-}_{k',t} + \mu^{se+}_{k',t} \geq 0, \forall k', \forall t < NT \tag{4.36}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial E_{k',t}} = \zeta_{k',t} - \xi^0_{k'} - \mu^{se-}_{k',t} + \mu^{se+}_{k',t} \geq 0, \forall k', t = NT \tag{4.37}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial n_t} = \lambda^b_t\, u_t - \lambda^s_t\,(u_t - 1) - \lambda^l_t \geq 0, \forall t \tag{4.38}$$

$$\frac{\partial \mathcal{L}_{LL4}}{\partial u_t} = \lambda^b_t\, n_t - \lambda^s_t\, n_t + \mu^{u+}_t \geq 0, \forall t \tag{4.39}$$

$$\mu^{d+}_{i',t} \geq 0, \forall i', \forall t \tag{4.40}$$

$$\mu^{g+}_{j',t} \geq 0, \forall j', \forall t \tag{4.41}$$

$$\mu^{se-}_{k',t}, \mu^{se+}_{k',t}, \mu^{sc+}_{k',t}, \mu^{sd+}_{k',t} \geq 0, \forall k', \forall t \tag{4.42}$$

$$\mu^{u+}_t \geq 0, \forall t \tag{4.43}$$

Constraints (4.32)-(4.39) constitute dual constraints with respect to primal variables $d_{i',t}$, $g_{j',t}$, $s^c_{k',t}$, $s^d_{k',t}$, $E_{k',t}$ and n_t respectively, while constraints (4.40)-(4.43) express non-negativity of the relevant dual variables.

Primal-dual Formulation of Relaxed LL4 problem

However, the relaxation of the binary constraints (4.26) implies that the reformulated LL4 problem (4.27)-(4.29) does not generally produce the optimal solution of the original LL4 problem (4.15)-(4.25). Thus, in order to ensure that the reformulated LL4 problem produces a solution that minimally deviates from the solution of the original LL4 problem, the following primal-dual formulation is adopted:

$$\min_{\{P^{LL4},D^{LL4}\}} \text{DG} \equiv -(4.15) - (4.30) \tag{4.44}$$

subject to:

$$(\text{Relaxed LL4 primal constraints}) : (4.17) \text{ and } (4.19) - (4.25) \tag{4.45}$$

$$(\text{Relaxed LL4 dual constraints}) : (4.32) - (4.43) \tag{4.46}$$

$$(\text{Original LL4 binary constraint}) : (4.18) \tag{4.47}$$

Problem (4.44)-(4.47) minimizes the duality gap (DG) between the primal and dual objective function values of the relaxed LL4 problem, while enforcing the primal constraints (4.17), (4.19)-(4.25) and the dual constraints (4.32)-(4.43) of the relaxed LL4 problem, as well as the original binary constraints (4.18). It should be noted that although the binary constraints were previously relaxed to enable the derivation of the dual problem, they are enforced in (4.47) to ensure that the solution of this problem conforms to the physical reality of LEM to buy energy from the retailer ($u_t = 1$) or sell energy to the retailer ($u_t = 0$) at period t.

4.4.7 Single-level Optimization Model Reduction

Formulation (4.44)-(4.47) constitutes the reformulation of the original LL4 problem. However, the optimal pricing problem from the perspective of a strategic retailer we aim at solving in this chapter is the bi-level problem (4.1)-(4.25) which additionally contains the one UL problem and the other three LL continuous problems. In order to achieve this in a mathematically rigorous fashion, we convert this bi-level problem to a single-level problem by the following two steps, which are illustrated in Fig. 4.3.

KKT Optimality Conditions Associated with Convex LL1-3 Problems

As discussed in previous section 4.2, the bi-level model can be solved by replacing its convex LL problems with its equivalent KKT optimality conditions and subject to the UL problems, rendering an MPEC. To this end, we derive the KKT optimality conditions of LL1-3 problems in the first step:

1. KKT Optimality Conditions Associated with the LL1 Problem

   To obtain the KKT conditions associated with the LL1 problem, the corresponding Lagrangian function $\mathcal{L}_{LL1}$ below is required as below:

$$\mathcal{L}_{LL1} = \sum_{i,t} \lambda_t^b \, d_{i,t} - \sum_{i,t} (l_{i,t}^D \, d_{i,t} - q_{i,t}^D \, d_{i,t}^2) - \sum_{i,t} \mu_{i,t}^{d-} d_{i,t} + \sum_{i,t} \mu_{i,t}^{d+} (d_{i,t} - d_{i,t}^{max}) \quad (4.48)$$

   Considering the Lagrangian function $\mathcal{L}_{LL1}$ given by (4.48), the KKT first order optimality conditions of the LL1 problem are derived as follows:

$$\frac{\partial \mathcal{L}_{LL1}}{\partial d_{i,t}} = \lambda_t^b - (l_{i,t}^D - 2\, q_{i,t}^D \, d_{i,t}) - \mu_{i,t}^{d-} + \mu_{i,t}^{d+} = 0, \forall i, \forall t \quad (4.49)$$

$$0 \leq \mu_{i,t}^{d-} \perp d_{i,t} \geq 0, \forall i, \forall t \quad (4.50)$$

$$0 \leq \mu_{i,t}^{d+} \perp (d_{i,t}^{max} - d_{i,t}) \geq 0, \forall i, \forall t \quad (4.51)$$

   The structure of the KKT conditions (4.49)-(4.51) is explained below:

   (a) Equality (4.49) is obtained by differentiating the Lagrangian function $\mathcal{L}_{LL1}$ with respect to the primal variables $d_{i,t}$.

   (b) Complementarity conditions (4.50)-(4.51) are related to the inequality constraints (4.6).

2. KKT Optimality Conditions Associated with the LL2 Problem

   To obtain the KKT conditions associated with the LL2 problem, the corresponding Lagrangian function $\mathcal{L}_{LL2}$ below is required as below:

$$\mathcal{L}_{LL2} = \sum_{j,t} (l_j^G \, g_{j,t} + q_j^G \, g_{j,t}^2) - \sum_{j,t} \lambda_t^s \, g_{j,t} - \sum_{j,t} \mu_{j,t}^{g-} g_{j,t} + \sum_{j,t} \mu_{j,t}^{g+} (g_{j,t} - g_j^{max}) \quad (4.52)$$

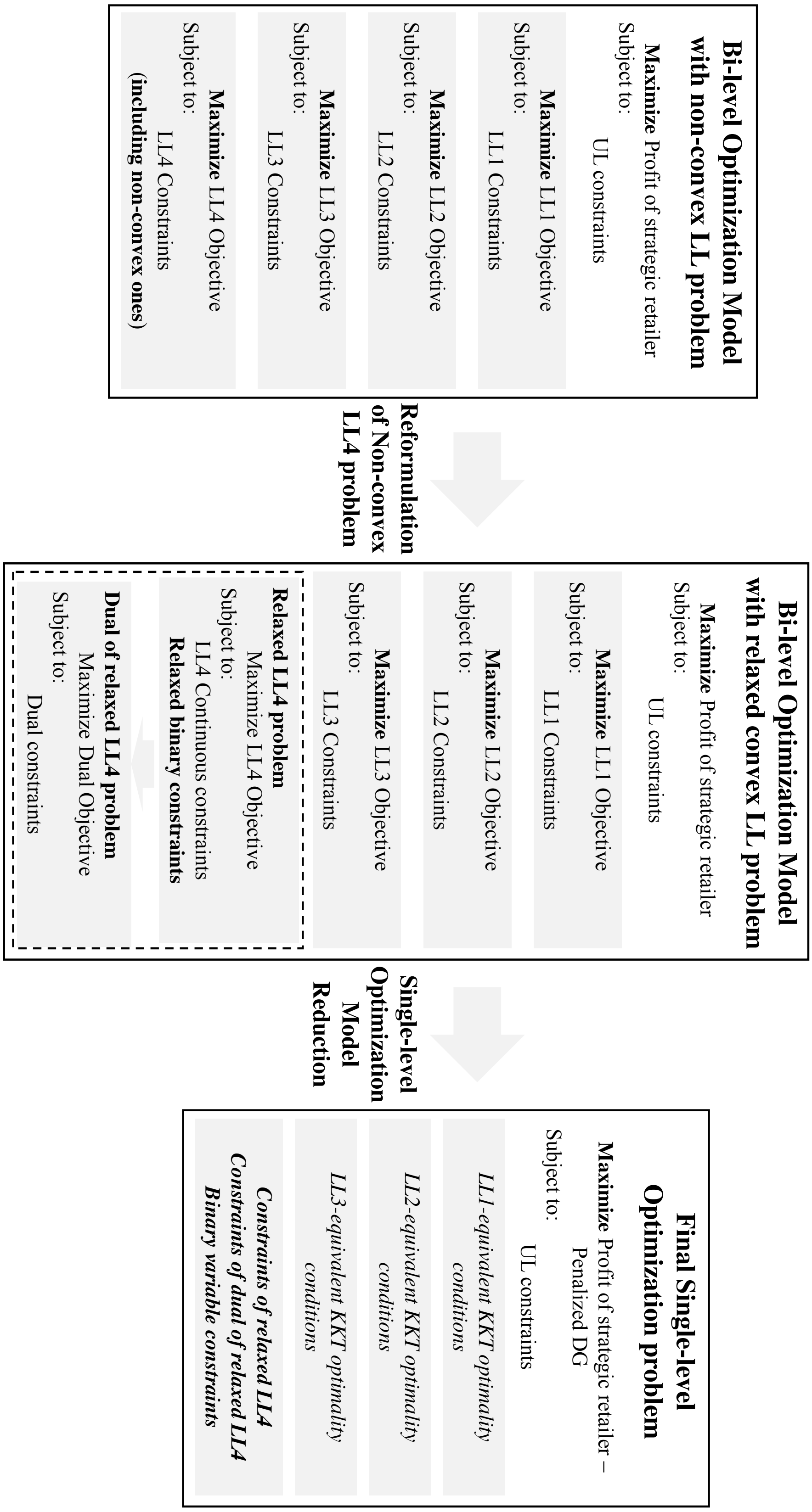

Fig. 4.3 Illustration of proposed modeling approach.

Considering the Lagrangian function $\mathcal{L}_{LL2}$ given by (4.52), the KKT first order optimality conditions of the LL2 problem are derived as follows:

$$\frac{\partial \mathcal{L}_{LL2}}{\partial g_{j,t}} = (l_j^G + 2\, q_j^G\, g_{j,t}) - \lambda_t^s - \mu_{j,t}^{g-} + \mu_{j,t}^{g+} = 0, \forall j, \forall t \tag{4.53}$$

$$0 \leq \mu_{j,t}^{g-} \perp g_{j,t} \geq 0, \forall j, \forall t \tag{4.54}$$

$$0 \leq \mu_{j,t}^{g+} \perp (g_j^{max} - g_{j,t}) \geq 0, \forall j, \forall t \tag{4.55}$$

The structure of the KKT conditions (4.53)-(4.55) is explained below:

(a) Equality (4.53) is obtained by differentiating the Lagrangian function $\mathcal{L}_{LL2}$ with respect to the primal variables $g_{j,t}$.

(b) Complementarity conditions (4.54)-(4.55) are related to the inequality constraints (4.8).

3. KKT Optimality Conditions Associated with the LL3 Problem

To obtain the KKT conditions associated with the LL3 problem, the corresponding Lagrangian function $\mathcal{L}_{LL3}$ below is required as below:

$$\begin{aligned}\mathcal{L}_{LL3} = &\sum_{k,t} \lambda_t^b\, s_{k,t}^c - \sum_{k,t} \lambda_t^s\, s_{k,t}^d + \sum_{k,t} \Big(\xi_{k,t}\,(E_{k,t} - E_{k,t-1} - s_{k,t}^c\, \eta_k^c + s_{k,t}^d/\eta_k^d)\Big) - \\ &\sum_{k,t} \mu_{k,t}^{se-}(E_{k,t} - E_k^{min}) + \sum_{k,t} \mu_{k,t}^{se+}(E_{k,t} - E_k^{max}) - \sum_{k,t} \mu_{k,t}^{sc-}\, s_{k,t}^c + \\ &\sum_{k,t} \mu_{k,t}^{sc+}(s_{k,t}^c - s_k^{max}) - \sum_{k,t} \mu_{k,t}^{sd+} s_{k,t}^d + \sum_{k,t} \mu_{k,t}^{sd+}(s_{k,t}^d - s_k^{max}) + \\ &\sum_{k} \xi_k^0 (E_k^0 - E_{k,NT})\end{aligned} \tag{4.56}$$

Considering the Lagrangian function $\mathcal{L}_{LL3}$ given by (4.56), the KKT first order optimality conditions of the LL3 problem are derived as follows:

$$\frac{\partial \mathcal{L}_{LL3}}{\partial s_{k,t}^c} = \lambda_t^b - \xi_{k,t}\, \eta_k^c - \mu_{k,t}^{sc-} + \mu_{k,t}^{sc+} = 0, \forall k, \forall t \tag{4.57}$$

$$\frac{\partial \mathcal{L}_{LL3}}{\partial s_{k,t}^d} = -\lambda_t^s + \xi_{k,t}/\eta_k^d - \mu_{k,t}^{sd-} + \mu_{k,t}^{sd+} = 0, \forall k, \forall t \tag{4.58}$$

$$\frac{\partial \mathcal{L}_{LL3}}{\partial E_{k,t}} = \xi_{k,t} - \xi_{k,t+1} - \mu_{k,t}^{se-} + \mu_{k,t}^{se+} = 0, \forall k, \forall t < NT \tag{4.59}$$

$$\frac{\partial \mathcal{L}_{LL3}}{\partial E_{k,t}} = \xi_{k,t} - \xi_k^0 - \mu_{k,t}^{se-} + \mu_{k,t}^{se+} = 0, \forall k, t = NT \tag{4.60}$$

$$0 \leq \mu_{k,t}^{se-} \perp (E_{k,t} - E_k^{min}) \geq 0, \forall k, \forall t \tag{4.61}$$

$$0 \leq \mu_{k,t}^{se+} \perp (E_k^{max} - E_{k,t}) \geq 0, \forall k, \forall t \tag{4.62}$$

$$0 \leq \mu_{k,t}^{sc-} \perp s_{k,t}^c \geq 0, \forall k, \forall t \tag{4.63}$$

$$0 \leq \mu_{k,t}^{sc+} \perp (s_k^{max} - s_{k,t}^c) \geq 0, \forall k, \forall t \tag{4.64}$$

$$0 \leq \mu_{k,t}^{sd-} \perp s_{k,t}^d \geq 0, \forall k, \forall t \tag{4.65}$$

$$0 \leq \mu_{k,t}^{sd+} \perp (s_k^{max} - s_{k,t}^d) \geq 0, \forall k, \forall t \tag{4.66}$$

The structure of the KKT conditions (4.57)-(4.66) is explained below:

(a) Equalities (4.57)-(4.60) are obtained by differentiating the Lagrangian function $\mathcal{L}_{LL3}$ with respect to the primal variables in the set of $\{s_{k,t}^c, s_{k,t}^d, E_{k,t}\}$.

(b) Complementarity conditions (4.61)-(4.66) are related to the inequality constraints (4.22)-(4.24).

Final Single-level Optimization Problem Formulation

In the second step, we combine the objective functions of the UL problem (4.1) and the reformulated LL4 problem (4.44)-(4.47) into a new designed objective function (4.67) and formulate it into a multi-objective optimization problem. In order to do this, we adopt the penalty function method [105] and penalize the DG by a positive constant W. By following this approach, the DG is indirectly suppressed (as prescribed by (4.44)), while still pursuing a higher retailer's profit (as prescribed by (4.1)). The value of the penalty constant W is selected by balancing the trade-off between the accuracy of the local energy market solution and the consideration of the retailer's profit. The impact of different values of W on the performance of the model and the selection of a suitable value are quantitatively analyzed in Section 4.5.3.

Enforcing all the constraints of the UL problem, the KKT optimality conditions of LL1-3 problems and the constraints of the reformulated LL4 problem. The resulting single-level problem is formulated as follows:

$$\max_{\{V\}} (4.1) - W * \text{DG} \tag{4.67}$$

where:

$$V = \{V^{UL}, P^{LL1}, P^{LL2}, P^{LL3}, P^{LL4}, D^{LL1}, D^{LL2}, D^{LL3}, D^{LL4}\} \tag{4.68}$$

subject to:

$$\text{(UL constraints)} : (4.2) - (4.4) \tag{4.69}$$

$$\text{(KKT optimality conditions of LL1)} : (4.49) - (4.51) \tag{4.70}$$

$$\text{(KKT optimality conditions of LL2)} : (4.53) - (4.55) \tag{4.71}$$

$$\text{(KKT optimality conditions of LL3)} : (4.57) - (4.66) \tag{4.72}$$

$$\text{(Primal-dual constraints of relaxed LL4)} : (4.45) - (4.47) \tag{4.73}$$

At this point, it should be stressed that this single-level optimization model is merely a tool used by the strtaegic retailer to determine its strategic making-decisions $\{\lambda_t^b, \lambda_t^s, w_t\}$ and is not generally fully accurate, as the proposed reformulation of the LL4 problem implies that it does not generally produce the optimal solution of the original LL4 problem. As a result, the profit (4.1) determined by the solution of model (4.67)-(4.73) is the profit estimated by the strategic retailer based on the proposed model (we will refer to it as *estimated profit* in the remainder) and is not generally equal to its *actual profit*, which is ultimately determined after inputting the values of optimal set $\{\lambda_t^b, \lambda_t^s, w_t\}$ obtained from model (4.67)-(4.73) to the actual response models (4.5)-(4.14) and local energy market operation (4.15)-(4.25).

4.4.8 MIQP

Furthermore, it should be noted that the resulting single-level optimization formulation is characterized by two types of non-convexity and thus any solution obtained by commercial solvers is not guaranteed to be globally optimal. Therefore, we aim at transforming this single-level optimization problem into a mixed-integer quadratic programming (MIQP), which can be efficiently solved to global optimality using commercial branch-and-cut solvers.

Linearizing Objective Function

The first one involves the bilinear terms in the objective function (4.67) of single-level problem, more specifically for (4.1) since the DG is a linear expression. We divided (4.1) into three parts: the business from individual response O^{ID}, the business in local energy market O^{LEM}, and the business in wholesale market O^{W}:

$$O^{ID} = \sum_t \lambda_t^b \Big(\sum_i d_{i,t} + \sum_k s_{k,t}^c \Big) - \sum_t \lambda_t^s \Big(\sum_j g_{j,t} + \sum_k s_{k,t}^d \Big) \tag{4.74}$$

$$O^{LEM} = \sum_t \lambda_t^b u_t n_t - \sum_t \lambda_t^s (u_t - 1) n_t \tag{4.75}$$

$$O^{W} = \sum_t \lambda_t^w w_t \tag{4.76}$$

Following the same steps in Chapter 3, the business from the individual response O^{ID} can be linearized by exploiting the strong duality theorem and some of the KKT optimiality conditions of LL1-3 problems. First of all, multiplying both sides of (4.49) by $d_{i,t}$, summing for every $i \in I$, $t \in T$ and rearranging some terms we get:

$$\sum_{i,t} \lambda_t^b \, d_{i,t} = \sum_{i,t} (l_{i,t}^D \, d_{i,t} - 2 \, q_{i,t}^D \, d_{i,t}^2) - \sum_{i,t} \mu_{i,t}^{d-} \, d_{i,t} + \sum_{i,t} \mu_{i,t}^{d+} \, d_{i,t} \tag{4.77}$$

For complementarity conditions (4.50) and (4.51):

$$\sum_{i,t} \mu_{i,t}^{d-} \, d_{i,t} = 0 \tag{4.78}$$

$$\sum_{i,t} \mu_{i,t}^{d+} \, d_{i,t} = \sum_{i,t} \mu_{i,t}^{d+} \, d_{i,t}^{max} \tag{4.79}$$

By substituting (4.78) and (4.79) into (4.77), renders the equality below:

$$\sum_{i,t} \lambda_t^b \, d_{i,t} = \sum_{i,t} (l_{i,t}^D \, d_{i,t} - 2 \, q_{i,t}^D \, d_{i,t}^2) + \sum_{i,t} \mu_{i,t}^{d+} \, d_{i,t}^{max} \tag{4.80}$$

Secondly, multiplying both sides of (4.53) by $g_{j,t}$, summing for every $j \in J$, $t \in T$ and rearranging some terms we get:

$$\sum_{j,t} \lambda_t^s \, g_{j,t} = \sum_{j,t} (l_j^G \, g_{j,t} + 2 \, q_j^G \, g_{j,t}^2) - \sum_{j,t} \mu_{j,t}^{g-} \, g_{j,t} + \sum_{j,t} \mu_{j,t}^{g+} \, g_{j,t} \tag{4.81}$$

For complementarity conditions (4.54) and (4.55):

$$\sum_{j,t} \mu_{j,t}^{g-} \, g_{j,t} = 0 \tag{4.82}$$

$$\sum_{j,t} \mu_{j,t}^{g+} \, g_{j,t} = \sum_{j,t} \mu_{j,t}^{g+} \, g_{j,t}^{max} \tag{4.83}$$

By substituting (4.78) and (4.83) into (4.81), renders the equality below:

$$\sum_{j,t} \lambda_t^s \, g_{j,t} = \sum_{j,t} (l_j^G \, g_{j,t} + 2 \, q_j^G \, g_{j,t}^2) + \sum_{j,t} \mu_{j,t}^{g+} \, g_j^{max} \tag{4.84}$$

Finally, We consider the LL3 problem is also a convex optimization problem, thus we can directly get the strong duality equality through the use of primal-dual transformation:

$$\sum_{k,t} \lambda_t^b \, s_{k,t}^c - \sum_{k,t} \lambda_t^s \, s_{k,t}^d = - \sum_{k,(t=1)} \xi_{k,t} \, E_k^0 + \sum_k \xi_k^0 \, E_k^0 + \sum_{k,t} \mu_{k,t}^{se-} \, E_k^{min} - \sum_{k,t} \mu_{k,t}^{se+} \, E_k^{max} - \sum_{k,t} \mu_{k,t}^{sc+} \, S_k^{max} - \sum_{k,t} \mu_{k,t}^{sd+} \, S_k^{max} \tag{4.85}$$

By combining (4.80), (4.84) and (4.4.8), renders the equality below:

$$O^{ID} = \sum_{i,t} (l_{i,t}^D \, d_{i,t} - 2 \, q_{i,t}^D \, d_{i,t}^2) + \sum_{i,t} \mu_{i,t}^{d+} \, d_{i,t}^{max} - \sum_{j,t} (l_j^G \, g_{j,t} + 2 \, q_j^G \, g_{j,t}^2) - \sum_{j,t} \mu_{j,t}^{g+} \, g_j^{max} - \sum_{k,(t=1)} \xi_{k,t} \, E_k^0 + \sum_k \xi_k^0 \, E_k^0 + \sum_{k,t} \mu_{k,t}^{se-} \, E_k^{min} - \sum_{k,t} \mu_{k,t}^{se+} \, E_k^{max} - \sum_{k,t} \mu_{k,t}^{sc+} \, S_k^{max} - \sum_{k,t} \mu_{k,t}^{sd+} \, S_k^{max} \tag{4.86}$$

However, the above techniques are not applicable to O^{LEM}, since its corresponding LL4 problem is non-convex and therefore strong duality does not hold. In order to address this challenge, the linearization of mixed-integer products [106] and the binary expansion approach [107, 108] are employed. The first step is introducing a new continuous variable $x_t^{net} = u_t n_t$ to linearize the product of the binary variable u_t and the continuous variable n_t, and formulate the $u_t n_t$ into the below mix-integer linear constraint (4.87):

$$M^{net} - (1 - u_t) M^{net} \leq x_t^{net} \leq u_t M^{net} \tag{4.87}$$

where M^{net} is the big-M values to bound the continuous variables n_t and is selected as tightly as possible. The second step is letting $\{\lambda_t^b, \, \lambda_t^s, \, l = 1, 2, ..., L\}$ be a set of discrete values in the range $[0, \lambda^{max}]$ as prescribed by the physical bounds (4.2)-(4.3). Then, the variables λ_t^b, λ_t^s can be expressed as the following sum of binary variables:

$$\lambda_t^b = \sum_{n=0}^{\log_2^L - 1} 2^n \triangle^b \, y_{t,n}^b, \forall t \tag{4.88}$$

$$\lambda_t^s = \sum_{n=0}^{\log_2^L - 1} 2^n \triangle^s y_{t,n}^s, \forall t \tag{4.89}$$

where $\triangle^b = \triangle^s = \frac{\lambda^{max}}{L-1}$ and $y_{t,n}^b, y_{t,n}^s$ are the auxiliary binary variables. Multiplying both sides of (4.88) and (4.89) by x_t^{net}, and both sides of (4.89) by n_t, summing for every t, and defining three dummy variables $z_{t,n}^b, z_{t,n}^s, z_{t,n}^{s0}$, results in:

$$\sum_t \lambda_t^b x_t^{net} = \sum_{t,n} 2^n \triangle^b z_{t,n}^b \tag{4.90}$$

$$\sum_t \lambda_t^s x_t^{net} = \sum_{t,n} 2^n \triangle^s z_{t,n}^s \tag{4.91}$$

$$\sum_t \lambda_t^s n_t = \sum_{t,n} 2^n \triangle^b z_{t,n}^{s0} \tag{4.92}$$

$$z_{t,n}^b = y_{t,n}^b x_t^{net} \tag{4.93}$$

$$z_{t,n}^s = y_{t,n}^s x_t^{net} \tag{4.94}$$

$$z_{t,n}^{s0} = y_{t,n}^s n_t \tag{4.95}$$

Therefore, the bilinear terms $\sum_t \lambda_t^b x_t^{net}$, $\sum_t \lambda_t^s x_t^{net}$ and $\sum_t \lambda_t^s n_t$ can be replaced by the expressions in the right side of (4.90), (4.91) and (4.92) which are all linear. The product of variables in (4.93), (4.94) and (4.95) can be transformed into the following equivalent mixed-integer linear constraints:

$$M^b - (1 - y_{t,n}^b)M^b \leq z_{t,n}^b \leq y_{t,n}^b M^b, \forall t, \forall n \tag{4.96}$$

$$M^s - (1 - y_{t,n}^s)M^s \leq z_{t,n}^s \leq y_{t,n}^s M^s, \forall t, \forall n \tag{4.97}$$

$$M^s - (1 - y_{t,n}^s)M^s \leq z_{t,n}^{s0} \leq y_{t,n}^s M^s, \forall t, \forall n \tag{4.98}$$

where M^b, M^s are two positive constants that are large enough for (4.96)-(4.97) to hold when $y_{t,n}^b = y_{t,n}^s = 0$ and $y_{t,n}^b = y_{t,n}^s = 1$, respectively.

Based on the above derivations, the business in the local energy market O^{LEM} can be expressed by the following expression, and subject to the above mixed-integer linear constraints (4.96)-(4.98):

$$O^{LEM} = \sum_{t,n} 2^n \triangle^b z_{t,n}^b - \sum_{t,n} 2^n \triangle^s z_{t,n}^s + \sum_{t,n} 2^n \triangle^b z_{t,n}^{s0} \tag{4.99}$$

Linearizing complementarity conditions

The second one is the bilinear terms in the complementarity conditions (4.50)-(4.51), (4.54)-(4.55), and (4.61)-(4.66), which can be expressed in the generic form $0 \leq \mu \perp p \geq 0$, with μ and p representing generic dual and primal terms respectively. The linearization approach proposed in [92] replaces each of these conditions with the set of mixed-integer linear conditions $\mu \geq 0$, $p \geq 0$, $\mu \leq \omega M^{\mu}$, $p \leq (1-\omega)M^{p}$, where ω is an auxiliary binary variable, while M^{μ} and M^{p} are large positive constants.

The values of the parameters M^{μ} and M^{p} should be suitably selected in order to achieve not only accurate but also computationally efficient solution of the MIQP. Specifically, M^{μ} and M^{p} should be large enough in order to avoid imposing additional upper bounds on the decision variables and thus resulting in an inaccurate solution of the MIQP. On the other hand, extremely large values should be avoided as they hinder the convergence of branch-and-cut solvers and result in large computational times [91]. Suitable values of the parameters M^{p} corresponding to primal terms can be more easily determined based on the bounds of primal variables which correspond to explicit physical limits. For example, the parameter M^{p} corresponding to the primal term of the complementarity constraint (4.50) is set equal to the maximum demand limit $d_{i,t}^{max}$ which physically limits the primal variable $d_{i,t}$. Suitable selection of the parameters M^{μ} corresponding to dual terms is more challenging since the dual variables do not exhibit explicit physical limits. In this context, the heuristic approach presented in [91] has been employed to tune parameters M^{μ}.

Final MIQP Formulation

Considering the three linearization techniques presented above, the final single-level optimization problem (4.67)-(4.73) can be transformed into the MIQP problem given by (4.100)-(4.178). Where the set of decision variables of the MIQP formulation includes the set (4.68) as well as the auxiliary binary variables introduced for linearizing (4.50)-(4.51), (4.54)-(4.55), and (4.61)-(4.66).

$$\max_{V,Z,S} \text{Profit} - W * \text{DG} \tag{4.100}$$

where:

$$\begin{aligned}\text{Profit} = &\sum_{i,t}(l^D_{i,t}\, d_{i,t} - 2\, q^D_{i,t}\, d^2_{i,t}) + \sum_{i,t}\mu^{d+}_{i,t}\, d^{max}_{i,t} - \sum_{j,t}(l^G_j\, g_{j,t} + 2\, q^G_j\, g^2_{j,t}) - \\ &\sum_{j,t}\mu^{g+}_{j,t}\, g^{max}_j - \sum_{k,(t=1)}\xi_{k,t}\, E^0_k + \sum_k \xi^0_k\, E^0_k + \sum_{k,t}\mu^{se-}_{k,t}\, E^{min}_k - \\ &\sum_{k,t}\mu^{se+}_{k,t}\, E^{max}_k - \sum_{k,t}\mu^{sc+}_{k,t}\, S^{max}_k - \sum_{k,t}\mu^{sd+}_{k,t}\, S^{max}_k + \\ &\sum_{t,n} 2^n \triangle^b z^b_{t,n} - \sum_{t,n} 2^n \triangle^s z^s_{t,n} + \sum_{t,n} 2^n \triangle^b z^{s0}_{t,n} - \sum_t \lambda^w_t\, w_t \end{aligned} \tag{4.101}$$

$$\begin{aligned}\text{DG} = &-\sum_{i',t}(l^D_{i',t}\, d_{i',t} - q^D_{i',t}\, d^2_{i',t}) + \sum_{j',t}(l^G_{j'}\, g_{j',t} + q^G_{j'}\, g^2_{j',t}) + \sum_{t,n} 2^n \triangle^b z^b_{t,n} - \\ &\sum_{t,n} 2^n \triangle^s z^s_{t,n} + \sum_{t,n} 2^n \triangle^b z^{s0}_{t,n} + \sum_{i',t}\mu^{d+}_{i',t} d^{max}_{i',t} + \sum_{j',t}\mu^{g+}_{j',t}\, g^{max}_{j'} - \\ &\sum_{k',t}\mu^{se-}_{k',t}\, E^{min}_{k'} + \sum_{k',t}\mu^{se+}_{k',t}\, E^{max}_{k'} + \sum_{k',t}\mu^{sc+}_{k',t}\, s^{max}_{k'} + \sum_{k',t}\mu^{sd+}_{k',t}\, s^{max}_{k'} + \\ &\sum_{k',(t=1)}\xi_{k',t}\, E^0_{k'} + \sum_{k'}\xi^0_{k'}\, E^0_{k'} + \sum_t \mu^{u+}_t \end{aligned} \tag{4.102}$$

$$\begin{aligned} V = \{&\lambda^b_t, \lambda^s_t, w_t, d_{i,t}, g_{j,t}, s^c_{k,t}, s^d_{k,t}, E_{k,t}, d_{i',t}, g_{j',t}, s^c_{k',t}, s^d_{k',t}, E_{k',t}, u_t, n_t, \\ &\mu^{d-}_{i,t}, \mu^{d+}_{i,t}, \mu^{g-}_{j,t}, \mu^{g+}_{j,t}, \xi_{k,t}, \mu^{se-}_{k,t}, \mu^{se+}_{k,t}, \mu^{sc-}_{k,t}, \mu^{sc+}_{k,t}, \mu^{sd-}_{k,t}, \mu^{sd+}_{k,t}, \xi^0_k, \\ &\lambda^l_t,\ \mu^{d+}_{i',t},\ \mu^{g+}_{j',t},\ \xi_{k',t},\ \mu^{se-}_{k',t},\ \mu^{se+}_{k',t},\ \mu^{sc+}_{k',t},\ \mu^{sd+}_{k',t},\ \xi^0_{k'},\ \mu^{u+}_t\} \end{aligned} \tag{4.103}$$

$$Z = \{z^b_{t,n}, z^s_{t,n}, z^{s0}_{t,n}, y^b_{t,n}, y^s_{t,n}\} \tag{4.104}$$

$$S = \{\omega^{d-}_{i,t}, \omega^{d+}_{i,t}, \omega^{g-}_{j,t}, \omega^{g+}_{j,t}, \omega^{se-}_{k,t}, \omega^{se+}_{k,t}, \omega^{sc-}_{k,t}, \omega^{sc+}_{k,t}, \omega^{sd-}_{k,t}, \omega^{sd+}_{k,t},\} \tag{4.105}$$

subject to:
UL constraints:

$$\lambda^{min} \le \lambda^b_t \le \lambda^{max}, \forall t \tag{4.106}$$

$$\lambda^{min} \le \lambda^s_t \le \lambda^{max}, \forall t \tag{4.107}$$

$$\sum_i d_{i,t} - \sum_j g_{j,t} + \sum_k (s^c_{k,t} - s^d_{k,t}) + n_t = w_t, \forall t \tag{4.108}$$

KKT optimality conditions and linearized complementary conditions of the LL1 problem:

$$\lambda^b_t - (l^D_{i,t} - 2\, q^D_{i,t}\, d_{i,t}) - \mu^{d-}_{i,t} + \mu^{d+}_{i,t} = 0, \forall i, \forall t \tag{4.109}$$

$$\mu_{i,t}^{d-} \geq 0, \forall i, \forall t \tag{4.110}$$

$$d_{i,t} \geq 0, \forall i, \forall t \tag{4.111}$$

$$\mu_{i,t}^{d-} \leq \omega_{i,t}^{d-} \, M^{\mu}, \forall i, \forall t \tag{4.112}$$

$$d_{i,t} \leq (1 - \omega_{i,t}^{d-}) \, M^{P}, \forall i, \forall t \tag{4.113}$$

$$\mu_{i,t}^{d+} \geq 0, \forall i, \forall t \tag{4.114}$$

$$d_{i,t}^{max} - d_{i,t} \geq 0, \forall i, \forall t \tag{4.115}$$

$$\mu_{i,t}^{d+} \leq \omega_{i,t}^{d+} \, M^{\mu}, \forall i, \forall t \tag{4.116}$$

$$d_{i,t}^{max} - d_{i,t} \leq (1 - \omega_{i,t}^{d+}) \, M^{P}, \forall i, \forall t \tag{4.117}$$

KKT optimality conditions and linearized complementary conditions of the LL2 problem:

$$(l_j^G + 2 \, q_j^G \, g_{j,t}) - \lambda_t^s - \mu_{j,t}^{g-} + \mu_{j,t}^{g+} = 0, \forall j, \forall t \tag{4.118}$$

$$\mu_{j,t}^{g-} \geq 0, \forall j, \forall t \tag{4.119}$$

$$g_{j,t} \geq 0, \forall j, \forall t \tag{4.120}$$

$$\mu_{j,t}^{g-} \leq \omega_{j,t}^{g-} \, M^{\mu}, \forall j, \forall t \tag{4.121}$$

$$g_{j,t} \leq (1 - \omega_{j,t}^{g-}) \, M^{P}, \forall j, \forall t \tag{4.122}$$

$$\mu_{j,t}^{g+} \geq 0, \forall j, \forall t \tag{4.123}$$

$$g_j^{max} - g_{j,t} \geq 0, \forall j, \forall t \tag{4.124}$$

$$\mu_{j,t}^{g+} \leq \omega_{j,t}^{g+} \, M^{\mu}, \forall j, \forall t \tag{4.125}$$

$$g_j^{max} - g_{j,t} \leq (1 - \omega_{j,t}^{g+}) \, M^{P}, \forall j, \forall t \tag{4.126}$$

KKT optimality conditions and linearized complementary conditions of the LL3 problem:

$$\lambda_t^b - \xi_{k,t} \, \eta_k^c - \mu_{k,t}^{sc-} + \mu_{k,t}^{sc+} = 0, \forall k, \forall t \tag{4.127}$$

$$- \lambda_t^s + \xi_{k,t} / \eta_k^d - \mu_{k,t}^{sd-} + \mu_{k,t}^{sd+} = 0, \forall k, \forall t \tag{4.128}$$

$$\xi_{k,t} - \xi_{k,t+1} - \mu_{k,t}^{se-} + \mu_{k,t}^{se+} = 0, \forall k, \forall t < NT \tag{4.129}$$

$$\xi_{k,t} - \xi_k^0 - \mu_{k,t}^{se-} + \mu_{k,t}^{se+} = 0, \forall k', t = NT \tag{4.130}$$

$$\mu_{k,t}^{se-} \geq 0, \forall k, \forall t \tag{4.131}$$

$$E_{k,t} - E_k^{min} \geq 0, \forall k, \forall t \tag{4.132}$$

$$\mu_{k,t}^{se-} \leq \omega_{k,t}^{se-} M^{\mu}, \forall k, \forall t \tag{4.133}$$

$$E_{k,t} - E_k^{min} \leq (1 - \omega_{k,t}^{se-}) M^P, \forall k, \forall t \tag{4.134}$$

$$\mu_{k,t}^{se+} \geq 0, \forall k, \forall t \tag{4.135}$$

$$E_k^{max} - E_{k,t} \geq 0, \forall k, \forall t \tag{4.136}$$

$$\mu_{k,t}^{se+} \leq \omega_{k,t}^{se+} M^{\mu}, \forall k, \forall t \tag{4.137}$$

$$E_k^{max} - E_{k,t} \leq (1 - \omega_{k,t}^{se+}) M^P, \forall k, \forall t \tag{4.138}$$

$$\mu_{k,t}^{sc-} \geq 0, \forall k, \forall t \tag{4.139}$$

$$s_{k,t}^{c} \geq 0, \forall k, \forall t \tag{4.140}$$

$$\mu_{k,t}^{sc-} \leq \omega_{k,t}^{sc-} M^{\mu}, \forall k, \forall t \tag{4.141}$$

$$s_{k,t}^{c} \leq (1 - \omega_{k,t}^{sc-}) M^P, \forall k, \forall t \tag{4.142}$$

$$\mu_{k,t}^{sc+} \geq 0, \forall k, \forall t \tag{4.143}$$

$$s_k^{max} - s_{k,t}^{c} \geq 0, \forall k, \forall t \tag{4.144}$$

$$\mu_{k,t}^{sc+} \leq \omega_{k,t}^{sc+} M^{\mu}, \forall k, \forall t \tag{4.145}$$

$$s_k^{max} - s_{k,t}^{c} \leq (1 - \omega_{k,t}^{sc+}) M^P, \forall k, \forall t \tag{4.146}$$

$$\mu_{k,t}^{sd-} \geq 0, \forall k, \forall t \tag{4.147}$$

$$s_{k,t}^{d} \geq 0, \forall k, \forall t \tag{4.148}$$

$$\mu_{k,t}^{sd-} \leq \omega_{k,t}^{sd-} M^{\mu}, \forall k, \forall t \tag{4.149}$$

$$s_{k,t}^{d} \leq (1 - \omega_{k,t}^{sd-}) M^P, \forall k, \forall t \tag{4.150}$$

$$\mu_{k,t}^{sd+} \geq 0, \forall k, \forall t \tag{4.151}$$

$$s_k^{max} - s_{k,t}^{d} \geq 0, \forall k, \forall t \tag{4.152}$$

$$\mu_{k,t}^{sd+} \leq \omega_{k,t}^{sd+} M^{\mu}, \forall k, \forall t \tag{4.153}$$

$$s_k^{max} - s_{k,t}^{d} \leq (1 - \omega_{k,t}^{sd+}) M^P, \forall k, \forall t \tag{4.154}$$

Primal-dual constraints of relaxed LL4 problem:

$$\sum_{i'} d_{i',t} - \sum_{j'} g_{j',t} + \sum_{k'} (s^c_{k',t} - s^d_{k',t}) = n_t, \forall t \tag{4.155}$$

$$0 \leq d_{i',t} \leq d^{max}_{i',t}, \forall i', \forall t \tag{4.156}$$

$$0 \leq g_{j',t} \leq g^{max}_{j'}, \forall j', \forall t \tag{4.157}$$

$$E_{k',t} = E_{k',t-1} + s^c_{k',t}\, \eta^c_{k'} - s^d_{k',t}/\eta^d_{k'}, \forall k', \forall t \tag{4.158}$$

$$E^{min}_{k'} \leq E_{k',t} \leq E^{max}_{k'}, \forall k', \forall t \tag{4.159}$$

$$0 \leq s^c_{k',t} \leq s^{max}_{k'}, \forall k', \forall t \tag{4.160}$$

$$0 \leq s^d_{k',t} \leq s^{max}_{k'}, \forall k', \forall t \tag{4.161}$$

$$E^0_{k'} = E_{k',NT}, \forall k' \tag{4.162}$$

$$\lambda^l_t - (l^D_{i',t} - 2\, q^D_{i',t}\, d_{i',t}) + \mu^{d+}_{i',t} \geq 0, \forall i', \forall t \tag{4.163}$$

$$(l^G_{j'} - 2\, q^G_{j'} g_{j',t}) - \lambda^l_t + \mu^{g+}_{j',t} \geq 0, \forall j', \forall t \tag{4.164}$$

$$\lambda^l_t - \zeta_{k',t}\, \xi^c_{k'} + \mu^{sc+}_{k',t} \geq 0, \forall k', \forall t \tag{4.165}$$

$$-\lambda^l_t + \xi_{k',t}/\eta^d_{k'} + \mu^{sd+}_{k',t} \geq 0, \forall k', \forall t \tag{4.166}$$

$$\zeta_{k',t} - \zeta_{k',t+1} - \mu^{se-}_{k',t} + \mu^{se+}_{k',t} \geq 0, \forall k', \forall t < NT \tag{4.167}$$

$$\zeta_{k',t} - \xi^0_{k'} - \mu^{se-}_{k',t} + \mu^{se+}_{k',t} \geq 0, \forall k', t = NT \tag{4.168}$$

$$\lambda^b_t\, u_t - \lambda^s_t\, (u_t - 1) - \lambda^l_t \geq 0, \forall t \tag{4.169}$$

$$\lambda^b_t\, n_t - \lambda^s_t\, n_t + \mu^{u+}_t \geq 0, \forall t \tag{4.170}$$

$$\mu^{d+}_{i',t} \geq 0, \forall i', \forall t \tag{4.171}$$

$$\mu^{g+}_{j',t} \geq 0, \forall j', \forall t \tag{4.172}$$

$$\mu^{se-}_{k',t}, \mu^{se+}_{k',t}, \mu^{sc+}_{k',t}, \mu^{sd+}_{k',t} \geq 0, \forall k', \forall t \tag{4.173}$$

$$\mu^{u+}_t \geq 0, \forall t \tag{4.174}$$

$$u_t \in \{0, 1\}, \forall t \tag{4.175}$$

Equivalent mixed-integer linear constraints to linearize the products of binary and continuous variables:

$$M^b - (1 - y^b_{t,n})M^b \leq z^b_{t,n} \leq y^b_{t,n}M^b, \forall t, \forall n \tag{4.176}$$

$$M^s - (1 - y^s_{t,n})M^s \leq z^s_{t,n} \leq y^s_{t,n}M^s, \forall t, \forall n \tag{4.177}$$

$$M^s - (1 - y^s_{t,n})M^s \leq z^{s0}_{t,n} \leq y^s_{t,n}M^s, \forall t, \forall n \tag{4.178}$$

4.5 Case Studies

4.5.1 Test Data and Implementation

The examined case studies apply the model developed in Section 4.4 in the context of a single day with hourly resolution, i.e. T = $\{1, 2, .., 24\}$. The examined retailer serves a set of customers including 12 FC of three different types, 6 MG of three different types and 2 identical ES, which are presented in Table 4.2. The total generation capacity of the 6 MG (Table 4.3) and the total power capacity of the 2 ES (Table 4.4) are assumed to be equal to 50% and 10% of the aggregate peak demand of the 12 FC (Fig. 4.4), respectively; these designing assumptions are made to ensure that a potential LEM among the retailer's customers can enable substantial amounts of energy trading between them, and examine conditions where the LEM can be both a buyer (during peak demand hours), but also a seller (during off-peak demand hours) of energy to the retailer. The assumed wholesale market prices across the examined day follow the pattern of a typical winter day in the UK, illustrated in Fig. 4.5. Finally, the maximum retail price is assumed $\lambda^{max} = 250$£/MWh.

Table 4.2 Energy use types and sizes of flexible consumers.

Consumer i	1-6	7-10	11-12
Type	Residential	Commercial	Industrial
Size (%)	32.88	37.00	30.12

The final MIQP model (Section 4.4.8) has been implemented and solved using the optimization software FICO™ Xpress [93] on a computer with a 6-core 3.50 GHz Intel(R) Xeon(R) E5-1650 v3 processor and 32 GB of RAM. It should be mentioned that this kind of relaxation and primal-dual reformulation approach is limited to the size of the problem. Given the examined case (12 FC, 6 MG, 2 ES) presented in this

Table 4.3 Operating parameters of micro-generators.

Micro-generator j	Type	l_j^G (£/MWh)	q_j^G (£/MW^{2}h)	g_j^G (MW)
1	Fuel cell	3	0.05	67.5
2	Fuel cell	3	0.05	67.5
3	Microturbine	8	0.09	40.5
4	Microturbine	8	0.09	40.5
5	Microturbine	13	0.14	27
6	Microturbine	13	0.14	27

Table 4.4 Operating parameters of energy storages.

Energy storage k	E_k^{max} (MWh)	E_k^{min} (MWh)	E_k^0 (MWh)	S_k^{max} (MW)	η_k^c, η_k^d (%)
1	108	21.6	27	27	90
2	108	21.6	27	27	90

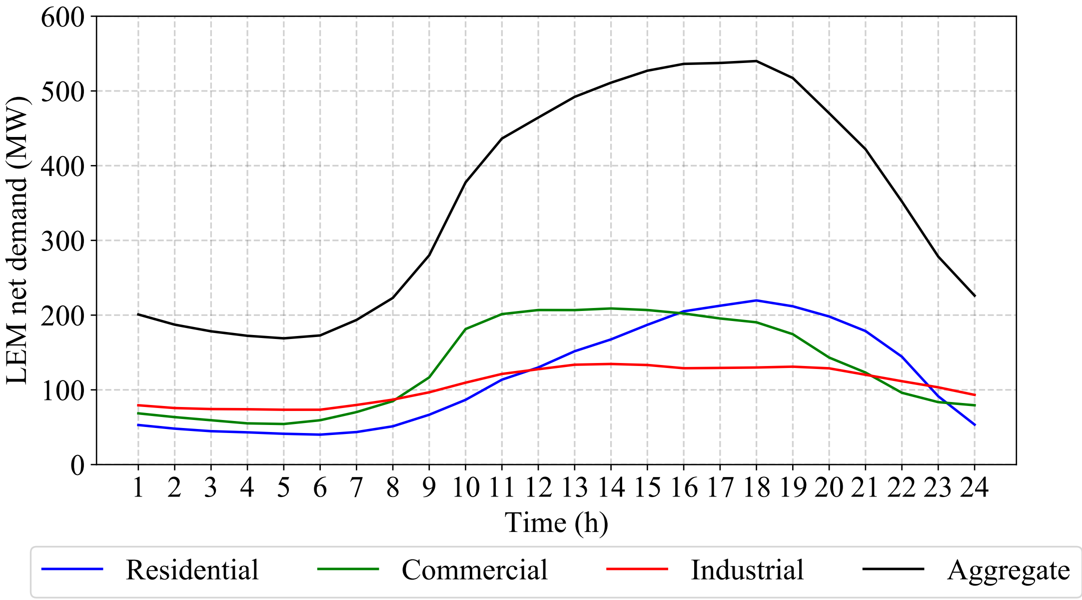

Fig. 4.4 Maximum demand of the aggregated consumers.

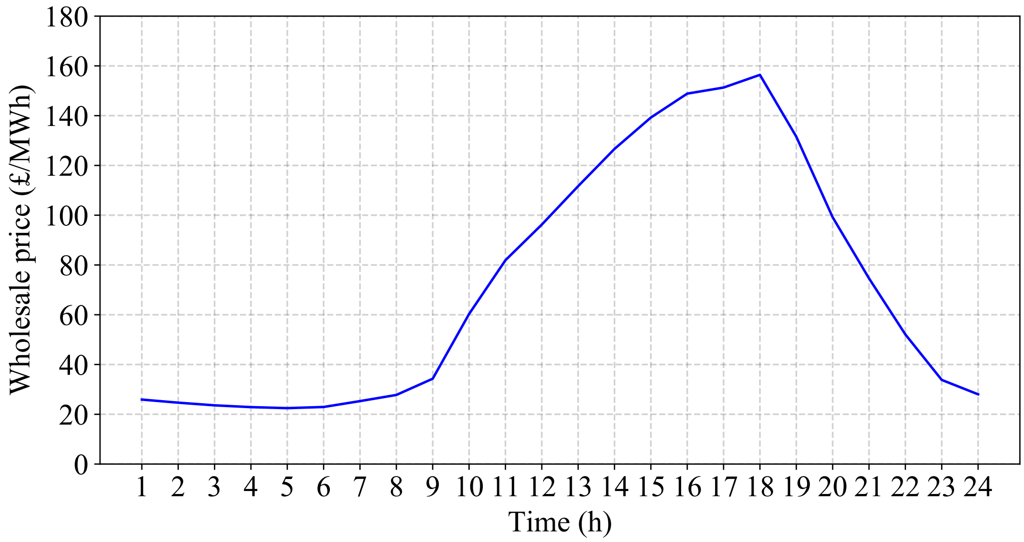

Fig. 4.5 Wholesale market price of the examined day.

work, the average computational time required for solving this MIQP across all the examined scenarios was around 62.8s.

4.5.2 Impact of Local Energy Market

The aim of the presented studies lies in analyzing the effects of introducing an LEM among the retailer's customers on both the retailer's and the customers' economic surplus. In this context, we examine three different scenarios regarding the composition of this LEM:

(i) No LEM: This constitutes the benchmark scenario where an LEM does not exist and all customers interact individually with the retailer.

(ii) Small LEM: Half of the retailer's customers (6 FC, 3 MG and 1 ES) participate in the LEM while the other half interacts individually with the retailer.

(iii) Large LEM: All customers participate in the LEM.

For each of these scenarios, Fig. 4.6 illustrates the hourly profiles of the net demand of the LEM n_t ($n_t > 0$ if the LEM buys energy from the retailer, $n_t < 0$ if the LEM sells energy to the retailer); Fig. 4.7 illustrates the hourly profiles of the aggregate charging / discharging power of the 2 ES (positive values indicate charging and negative values indicate discharging); Fig. 4.8 illustrates the hourly profiles of the total customers' demand served by the retailer ($\sum_i d_{i,t} + \sum_k s^c_{k,t} + u_t n_t$); and Fig. 4.9 illustrates the hourly profiles of the total customers' generation served by the retailer ($\sum_j g_{j,t} + \sum_k s^d_{k,t} + (u_t - 1)n_t$).

Furthermore, Fig. 4.10 illustrates the hourly profiles of the retail buy prices λ^b_t offered by the retailer (and the wholesale prices λ^w_t for comparison purposes); Fig. 4.11 illustrates the hourly profiles of the retail sell prices λ^s_t offered by the retailer (and the wholesale prices λ^w_t for comparison purposes); Fig. 4.12 illustrates the hourly profiles of the clearing prices of the LEM (in the two scenarios with LEM), which correspond to the dual variables of constraints (4.21) (and the buy and sell prices offered by the retailer in the No LEM scenario for comparison purposes); and Fig. 4.13 illustrates the hourly profiles of the comparison between the retail buy prices, retail sell prices and LEM clearing prices in the Small LEM scenario.

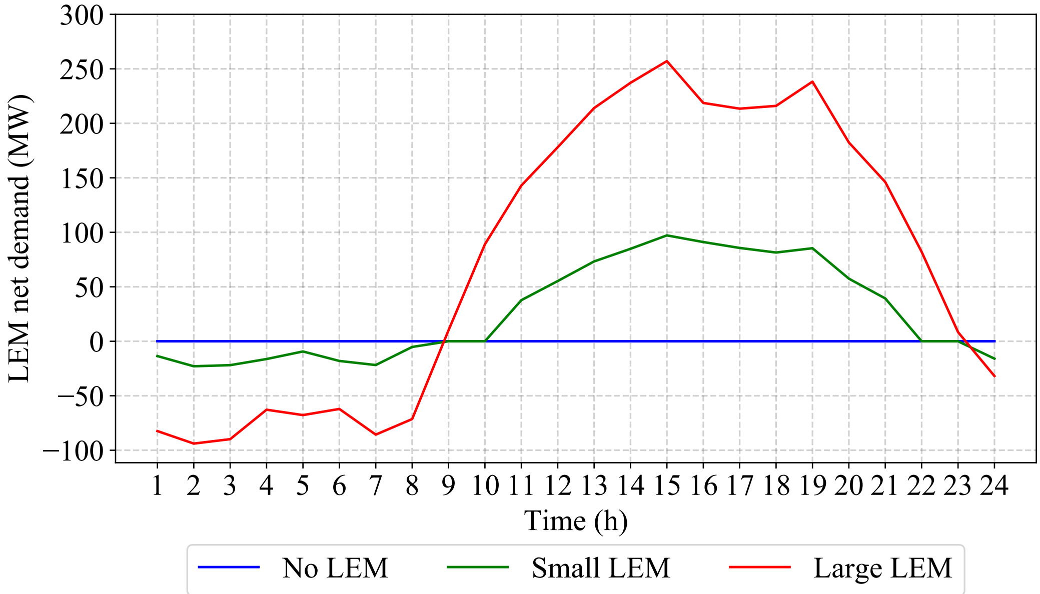

Fig. 4.6 Net demand of LEM for different scenarios.

In the benchmark scenario (No LEM scenario), both the total demand and the total generation served by the retailer exhibit the highest values across all hours (Fig. 4.8 and 4.9, respectively). This is because in the absence of an LEM, the only option for the customers to buy and sell energy is through the retailer. The strategic retailer exploits this dependency of its customers by offering very high buy prices (significantly higher than the wholesale prices) to demanding customers (Fig. 4.10) and very low sell prices (significantly lower than the wholesale prices) to generating customers (Fig. 4.11); please note that, for clarity purposes, the high wholesale prices during hours

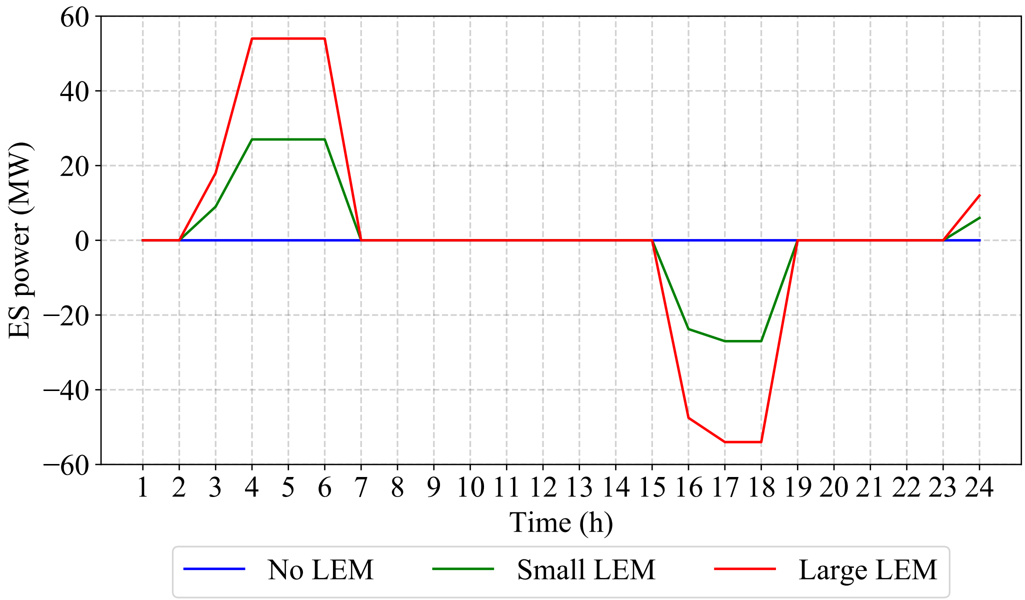

Fig. 4.7 Aggregare charging / discharging power of all ES for different scenarios.

9-22 are not presented in this figure) in order to maximize its overall profit (4.1). In other words, the strategic retailer exploits the customers by setting large differentials between buy and sell prices.

The offered buy prices are higher during peak hours, driven by a combination of the higher energy requirements of the retailer's consumers and the higher wholesale prices. The offered sell prices are also higher during peak hours, as the retailer strives to attract higher generation by its customers in order to reduce its net demand in the wholesale market which exhibits higher prices at these hours. However, even the lowest buy price (offered at hour 5) is higher than the highest sell price (offered at hours 10-21), rendering energy arbitrage non-profitable. As a result, the ES served by the retailer prefer to remain idle (neither charge nor discharge throughout the day) in order to avoid negative profits (Fig. 4.7).

When an LEM is introduced (Small LEM and Large LEM scenarios), the participating customers have the additional option (apart from trading energy with the retailer) to trade energy between them. Given that the retailer offers very high buy prices and very low sell prices in the No LEM scenario, the LEM participants choose to exercise this option and trade energy at intermediate prices (Fig. 4.12), which is mutually beneficial for all FC, MG and ES participants. As a result, the dependency of the LEM participants on the retailer is limited; both the total demand and the total generation served by the retailer are significantly reduced (Fig. 4.8 and 4.9, respectively). It is

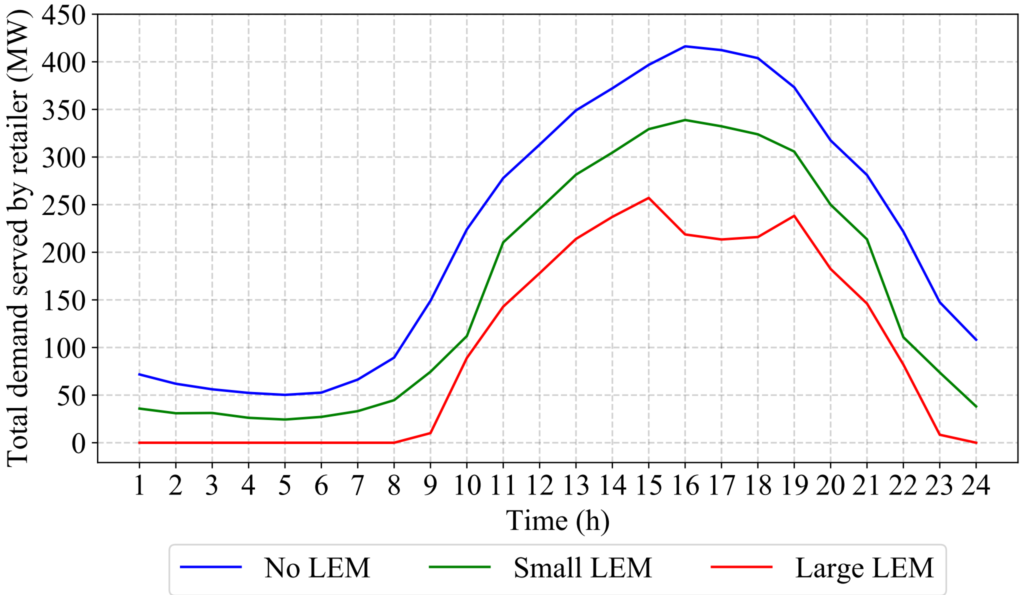

Fig. 4.8 Total demand served by the retailer for different scenarios.

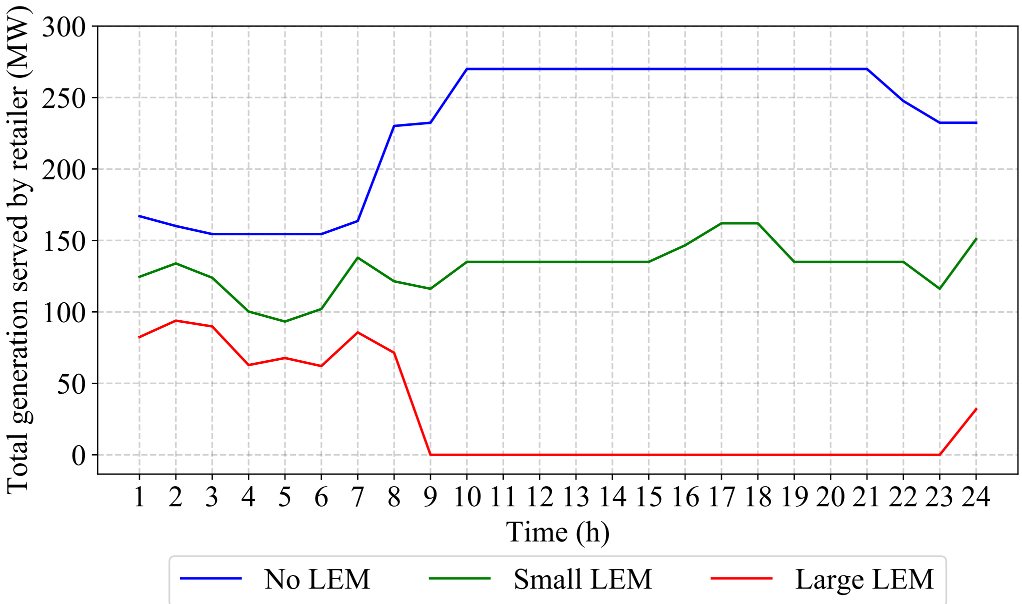

Fig. 4.9 Total generation served by the retailer for different scenarios.

worth noting however that although this dependency is reduced, it is not completely eliminated; the LEM still: i) buys energy from the retailer at peak hours, since its available generation and storage capacity is not sufficient to cover the high demand requirements during these hours, and ii) sells energy to the retailer at off-peak hours, since its available generation capacity is higher than the low demand requirements during these hours (Fig. 4.6).

Regarding the ES customers in particular, it is worth stressing that in these two scenarios, in contrast with the No LEM scenario, the ES participating in the LEM do not remain idle and carry out charging / discharging actions (Fig. 4.7), since such actions are beneficial for both: i) the ES participants themselves, given that they now face the LEM clearing prices which, in contrast to the retailer's differentiated buy and sell prices, render energy arbitrage profitable, as they are higher during peak hours and lower during off-peak hours (Fig. 4.12), and ii) the LEM as a whole, given that these actions further limit the dependency of the LEM on the retailer, as less demand is served by the retailer during peak hours (due to ES discharging) and less generation is served by the retailer during off-peak hours (due to ES discharging). In the same vein, it is noted that in the Small LEM scenario, the charging / discharging actions illustrated in Fig. 4.7 correspond to the ES participating in the LEM while the other ES still remains idle.

As expected, this effect of limiting the dependency of the customers on the retailer is enhanced when all customers participate in the LEM (Large LEM scenario) with respect to the scenario where only half of them participate (Small LEM scenario). In particular, during the off-peak hours 24-8, the demand served by the retailer becomes zero (Fig. 4.8), as the available generation capacity is able to cover the low demand requirements; on the other hand, during the peak hours 9-23, the generation served by the retailer becomes zero (Fig. 4.9), as the whole generation and storage capacity is utilized within the LEM to cover a part of the high demand requirements and limit as much as possible the dependency of the consumers on the retailer.

As a result of this effect of LEM in limiting the dependency of the customers on the retailer, the latter reduces substantially the offered buy prices across the majority of hours (Fig. 4.10) in order to attract higher demand from its demanding customers and compensate the reduction of its served demand caused by the LEM. In mathematical terms, this is due to the fact that the revenue of the retailer (first term of its overall profit (4.1)) depends on the summation of the buy price-demand products across all hours, i.e. the retailer needs to balance the trade-off between buy prices and customers' demand. In the Large LEM scenario and during the off-peak hours 24-8, given that

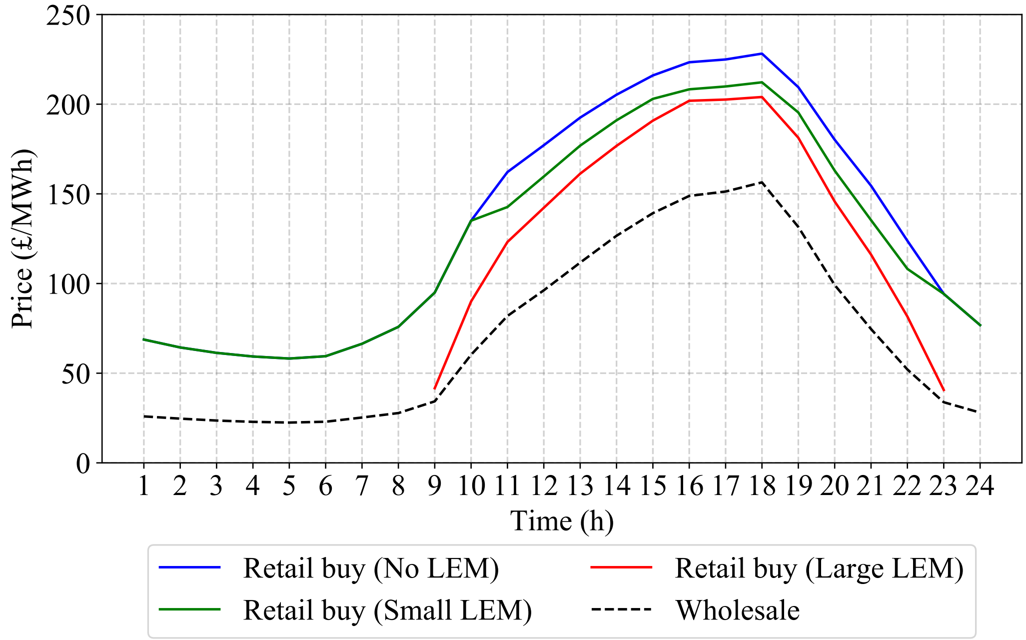

Fig. 4.10 Buy prices offered by the retailer for different scenarios.

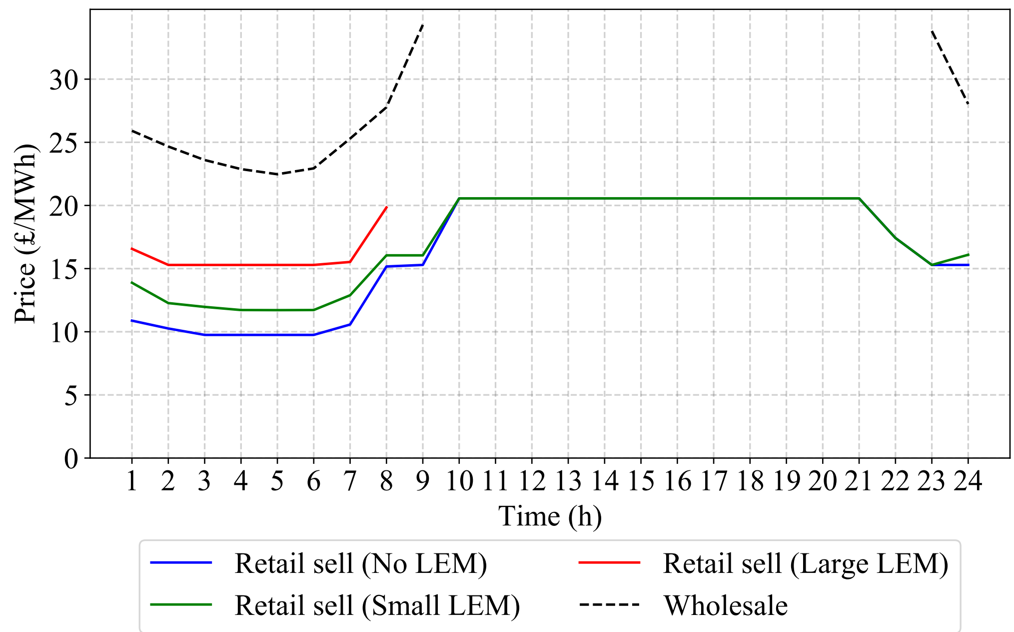

Fig. 4.11 Sell prices offered by the retailer for different scenarios.

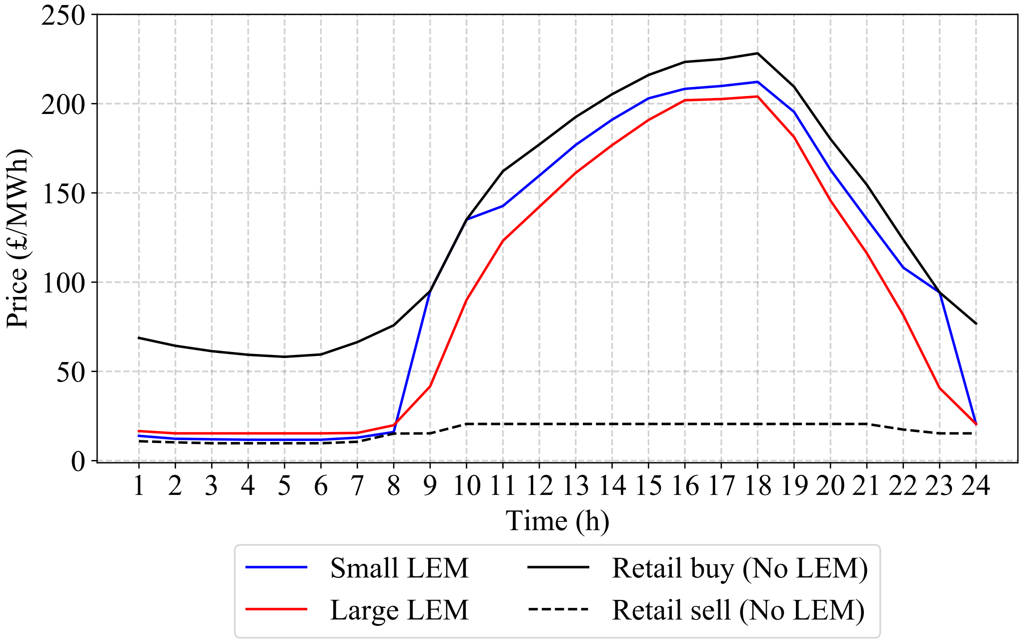

Fig. 4.12 LEM clearing prices for different scenarios.

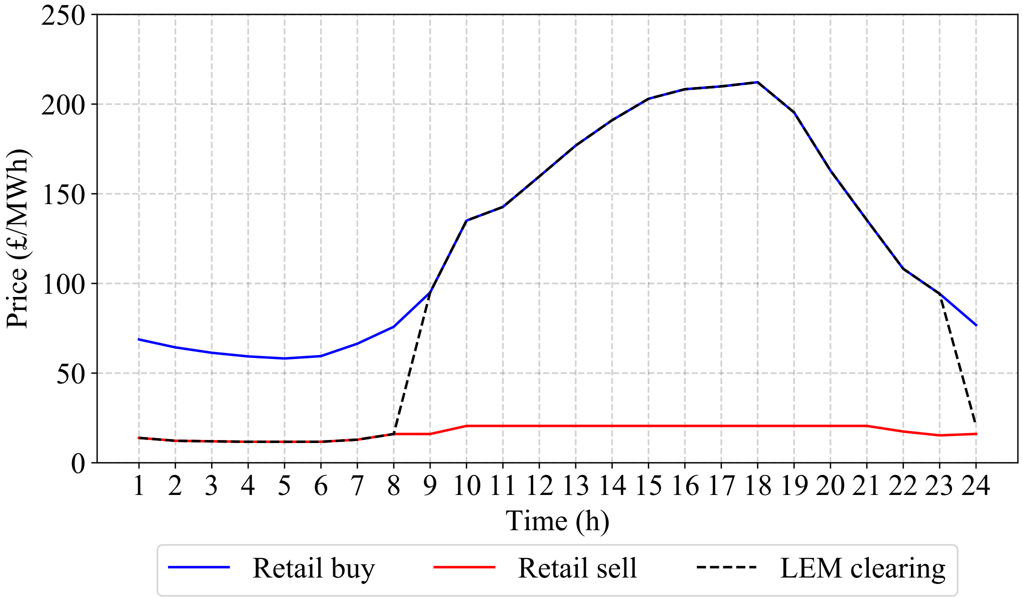

Fig. 4.13 Buy prices, sell prices and LEM clearing prices in the Small LEM scenario.

the demand served by the retailer is zero (Fig. 4.8), the offered buy prices have no physical impact (in mathematical terms they constitute free variables) and thus are omitted from Fig. 4.10.

In a similar spirit, the retailer increases substantially the offered sell prices across the majority of hours (Fig. 4.11) in order to attract higher generation from its generating customers and compensate the reduction of its served generation caused by the LEM. In mathematical terms, although this increase of the sell prices and customers' generation increases its cost at the retail side (second term of its overall profit (4.1)), it reduces its net demand w_t in the wholesale market (4.4) and therefore reduces its cost (during peak hours) and increases its revenue (during off-peak hours) at the wholesale side (third term of its overall profit (4.1)). In the Large LEM scenario and during the peak hours 9-23, given that the generation served by the retailer is zero (Fig. 4.9), the offered sell prices have no physical impact (in mathematical terms they constitute free variables) and thus are omitted from Fig. 4.11.

From a higher-level perspective, the above trends imply that the introduction of the LEM limits the retailer's strategic potential of exploiting the customers by setting large differentials between buy and sell prices. The offered buy prices are reduced and the offered sell prices are increased, i.e. they both move closer to the wholesale prices, implying that the retailer behaves more competitively.

It is also worth noting that in the Small LEM scenario and during off-peak hours, the LEM clearing prices are equal to the sell prices offered by the retailer (since the LEM sells its excess generation to the retailer), which are significantly lower than the buy prices offered at the same hours (Fig. 4.13); this implies that FC participating in the LEM gain significant benefits with respect to individual FC during these hours. In a similar vein, during peak hours, the LEM clearing prices are equal to the buy prices offered by the retailer (since the LEM buys its excess demand from the retailer), which are significantly higher than the sell prices offered at the same hours; this implies that MG participating in the LEM gain significant benefits with respect to individual MG during these hours.

Having analyzed the fundamental interactions between the retailer and the LEM through the demand / generation profiles of Fig. 4.6-4.9 and the price profiles of Fig. 4.10-4.13, the final part of this analysis lies in quantifying and analyzing the overall economic effects of the LEM on both the retailer and its customers. Starting from the former, Table I presents the total (daily) profit of the examined retailer and its three components (corresponding to the three terms of (4.1)), for each of the examined scenarios.

Table 4.5 Profit of Retailer for Different Scenarios.

Scenario	Revenue (£)	Retail cost (£)	Wholesale net cost (£)	Profit (£)
No LEM	913,313	96,288	93,645	723,380
Small LEM	645,532	61,635	174,032	409,866
Large LEM	402,109	10,522	271,815	119,772

We start our analysis from the retailer's revenue which, as demonstrated in Table 4.5, constitutes the most significant component of its overall profit. This revenue is greatly reduced in the Small LEM scenario (29% with respect to the No LEM scenario) and further reduced in the Large LEM scenario (56% with respect to the No LEM scenario). This trend is driven by the fact that the introduction of the LEM reduces both the total demand served by the retailer (Fig. 4.8) and the buy prices offered by the retailer (Fig. 4.10).

In a similar vein, the retailer's cost of buying energy from its generating customers is substantially reduced (36% in the Small LEM scenario and 89% in the Large LEM scenario). This trend is driven by the fact that the introduction of the LEM reduces significantly the total generation served by the retailer (Fig. 4.9), despite the relatively small increase of the sell prices offered by the retailer (Fig. 4.11).

Going further, the retailer's net cost in the wholesale market is significantly increased (86% in the Small LEM scenario and 190% in the Large LEM scenario). This trend is driven by the fact that the introduction of the LEM reduces the total generation served by the retailer to a higher extent than the reduction of the total demand served by the retailer (Fig. 4.8 and 4.9), especially during peak hours when the difference between the LEM clearing prices and the sell prices offered by the retailer are extremely high (Fig. 4.12). As a result, the retailer needs to buy more energy in the wholesale market, especially during peak periods which are characterized by high wholesale prices.

Overall, the introduction of the LEM reduces very significantly the retailer's total profit (43% in the Small LEM scenario and 83% in the Large LEM scenario), driven primarily by the reduction of its retail net revenue and secondarily by the increase of its wholesale net cost.

Moving our focus to the customers, Tables 4.6-4.8 present the total (daily) economic surplus of two identical FC, two identical MG and the two (identical) ES. The only difference between the two members of each of these sets of customers lies in the fact that, in the Small LEM scenario, the first member (the one corresponding to the second column of Tables 4.6-4.8) participates in the LEM while the second (the one corresponding to the third column of Tables 4.6-4.8) does not.

Table 4.6 Utility of Flexible Consumers for Different Scenarios.

Scenario	Utility of FC1 (£)	Utility of FC2 (£)
No LEM	32,890	32,890
Small LEM	55,220	46,279
Large LEM	70,327	70,327

Table 4.7 Profit of Micro-generators for Different Scenarios.

Scenario	Profit of MG1 (£)	Profit of MG2 (£)
No LEM	16,395	16,395
Small LEM	154,956	20,266
Large LEM	141,460	141,460

Table 4.8 Profit of Energy Storages for Different Scenarios.

Scenario	Profit of ES1 (£)	Profit of ES2 (£)
No LEM	0	0
Small LEM	14,991	0
Large LEM	14,276	14,276

The surplus of most customers is increased in the Small LEM scenario with respect to the No LEM scenario. This beneficial impact of the LEM is evident even for FC and MG not participating in the LEM, driven by the fact that the introduction of the LEM reduces the buy prices and increases the sell prices offered by the retailer (Fig. 4.10 and 4.11). The only exception to this beneficial impact is ES2; in a similar fashion with the response of the two ES in the No LEM scenario, ES2 remains idle and thus makes zero profit (Table 4.8), since the lowest buy price is still higher than the highest sell price, despite the above effects of the LEM (Fig. 4.13).

However, this beneficial impact is significantly higher for customers participating in the LEM, since they trade energy based on the LEM clearing prices. As previously discussed, during off-peak hours, these prices are substantially lower than the buy prices offered by the retailer, creating additional benefits for FC1 participating in the LEM with respect to the individual FC2 (Table 4.6). In a similar vein, during peak hours, the LEM clearing prices are significantly higher than the sell prices offered by the retailer, creating additional benefits for MG1 with respect to the individual MG2 (Table 3.1); it is worth noting that the surplus difference between MG1 and MG2 is significantly higher than the one between FC1 and FC2, due to the extremely high difference between the LEM clearing prices and the retailer's sell prices during peak hours. Considering ES1 participating in the LEM, in contrast with the individual ES2, it does not remain idle and carries charging / discharging actions which generate positive profits (Table 4.8), since it faces the LEM clearing prices which are higher during peak hours and lower during off-peak hours (Fig. 4.12).

In the Large LEM scenario, the utility of FC1 is further increased with respect to the Small LEM scenario, since the LEM clearing prices are reduced during peak hours (Fig. 4.12), driven by the participation of more generation and storage capacity in the LEM. On the other hand, this reduction of the LEM clearing prices during peak hours causes a reduction of the profit of MG1 and ES1 with respect to the Small LEM scenario. Finally, the surplus of FC2, MG2 and ES2 increases, since, in contrast with the Small LEM scenario, they now participate in the LEM and face the more favorable LEM clearing prices.

4.5.3 Effectiveness of Proposed Relaxation Approach

As discussed in Section 4.4.7, the proposed single-level optimization model is a multi-objective optimization problem, thus the effectiveness of the proposed model highly depends on the selection of W. In order to quantitatively demonstrate this, we select the large community scenario as the examined case, execute the proposed model for different values of W and compare its respective solutions in terms of: i) the DG determined by the proposed model; ii) the minimum DG obtained in (4.44)-(4.47) given the optimal value of $\lambda_t^b, \lambda_t^s, w_t$ determined by the proposed model; iii) the profit of retailer determined by the proposed model; and iv) the actual profit of retailer obtained through the actual response problems (4.5)-(4.25) after inputting the optimal value of λ_t^b, λ_t^s determined by the proposed model. The results of this analysis are included in Tables 4.9.

These results demonstrate that relatively small values of W (1-100) do not sufficiently penalize the DG in the objective function (4.100) and therefore the optimal DG is higher than the minimum DG. This means that the model is not as accurate as it could be in approximating the community response solution of the original LL4 problem. Intuitively, this implies a higher profit than the actual profit as the objective

Table 4.9 Performance of Proposed Approach for Different Values of Penality Value W.

W	Proposed DG (£)	Mini DG (£)	Proposed Profit (£)	Actual Profit (£)
1	6,887	5,689	125,012	108,843
10	5,839	4,641	123,066	110,041
100	3,593	3,144	120,820	116,628
1,000	1,647	1,647	120,071	119,772
10,000	1,198	1,198	117,077	118,425

of optimization concentrates on profit maximization. However, the optimal profit is far away from the actual profit, leading to the over-optimistic strategies. Driven by this reason, increasing values of W corrects the above problem, the optimal DG equals to the minimum DG above 1,000. However, a significant higher value of W results in another issue, retailer's profit becomes less important in the objective (4.100) and leads to the under-optimistic strategies. Eventually, the best value selection of W exists the highest actual profit with the zero difference between the optimal DG and the minimum DG. In this case, $W = 1,000$ (Table 4.9) is the one to be selected.

4.6 Conclusions

This chapter has explored for the first time the interactions between the operation of LEM with different types of participants (FC, MG and ES) and the strategic pricing decisions of incumbent electricity retailers, and has quantitatively analyzed the overall economic effects of LEM on both the retailer and its customers. In order to achieve that, this chapter has developed a novel multi-period bi-level optimization model, which captures the pricing decisions of a retailer in the upper level (UL) problem and the response of both independent customers and the LEM in the lower level (LL) problems. Since the LL problem representing the LEM is non-convex, the traditional MPEC approach is not applicable for solving the developed bi-level problem, and a new approach recently proposed by the authors is employed instead, which is based on the relaxation and primal-dual reformulation of the non-convex LL problem and the penalization of the associated duality gap.

The presented case studies have provided numerous new and valuable insights around the role and impact of LEM. First of all, the introduction of an LEM is shown to reduce the customers' energy dependency on the retailer, since they are able to trade energy among them at prices which lie between the retailer's high buy prices and low sell prices, which is mutually beneficial for all FC, MG and ES participants; regarding the latter, the LEM is shown to unlock their arbitrage potential and activate them in the market. As a consequence, the retailer's strategic potential of exploiting the customers through large differentials between buy and sell prices is limited, and the retailer strives to make its offered buy and sell prices more competitive in order to attract more demand and generation by its customers. As a result of these effects, the profit of the retailer is very significantly reduced, while the customers enjoy significant economic benefits. Although this beneficial impact of LEM is significantly higher for customers participating in the LEM, it is also substantial for non-participating customers, due to the above effects of the LEM on the retailer's offered prices.

Chapter 5

Strategic Pricing for Electric Vehicles with Discrete Charging

5.1 Introduction

Environmental and energy security concerns have driven governments worldwide to take significant initiatives towards the decarbonization of both generation and demand sides of energy systems [1]. However, these decarbonization initiatives introduce significant techno-economic challenges to the operation and development of electricity systems. At the generation side, the decarbonization agenda involves the large-scale integration of renewable generation, which is however inherently characterized by high variability and limited controllability, challenging the cost-efficient balancing of the electricity system. At the demand side, the decarbonization agenda involves the electrification of certain sectors, with the electrification of the transport sector through the large-scale integration of electric vehicles (EV) being one of the key priorities. However, due to the natural energy intensity and temporal demand patterns of transportation vehicles, this paradigm change is accompanied by a significant increase of demand peaks, driving capital-intensive generation and network investments [94].

In this setting, flexible demand technologies, enabling redistribution of electricity demand in time, have recently attracted great interest, since they exhibit the potential to support system balancing and limit demand peaks, thus improving significantly the cost-effectiveness of low-carbon electricity systems [94, 2, 6]. Amongst such flexible demand technologies, EV exhibit an outstanding flexibility potential due to their inherent ability to store electrical energy in their batteries, their low energy consumption requirements with respect to the significant capacity of their batteries, and the Vehicle-to-Grid (V2G) capability which enables EV to inject stored energy back to the grid [109–111].

Beyond the above decarbonization initiatives however, governments worldwide have also taken significant initiatives towards the deregulation of the electricity industry, involving unbundling of vertically integrated monopoly utilities and the introduction of competition in both generation and retail sectors [7]. In this deregulated environment, the realization of the EV flexibility potential needs to be driven by their full and suitable integration in electricity markets. One of the most promising mechanisms to achieve such market integration is EV aggregation [112], given that individual EV do not have sufficient capacity to participate independently in the wholesale electricity market. An EV aggregator (which is a role that can also be taken in practice by electricity suppliers) represents a large number of EV in the wholesale market and coordinate their operation according to the market conditions and the EV operating characteristics (e.g. travelling patterns and battery's / charger's operating parameters) to maximize its overall profit.

Since centralized, direct coordination approaches suffer from communication, computational and privacy limitations, price-based coordination approaches have lately attracted significant interest [113]. The EV aggregator offers certain retail electricity prices to the EV which then independently determine their optimal charging / discharging responses by solving their own cost minimization problems. The prices employed by the aggregator are usually time-specific in order to activate the EV flexibility and exploit favorable market conditions. In this context, the EV aggregator requires a suitable mechanism to design effective time-specific prices to maximize its overall profit, accounting for the response patterns of its EV.

5.2 Literature Review

Bi-level optimization constitutes the most widely employed methodological framework in the existing literature for addressing this problem [31, 34, 114, 115]. The popularity of this methodology lies in its ability to capture the interactions between the pricing optimization problem of the EV aggregator (modeled as the upper level problem - UL) and the response optimization problems of the EV (modeled as the lower level problems - LL). All relevant existing works [31, 34, 114, 115] have solved this bi-level optimization problem by converting it to a single-level *Mathematical Program with*

Equilibrium Constraints (MPEC), through the replacement of the LL problems by their equivalent *Karush-Kuhn-Tucker* (KKT) optimality conditions.

Nevertheless, this solution approach exhibits two fundamental limitations. First of all, it implicitly assumes that the aggregator (solving the final MPEC) has perfect knowledge of the EV operating characteristics (which constitute parameters of the LL problems) and imperfect knowledge of the EV travelling patterns (which constitute uncertainties of the LL problems); such an assumption is not generally realistic, particularly when considering that the current penetration of EV is limited and thus existing work on the systematic characterization of their operating characteristics (especially their travelling patterns) is far from comprehensive [116]. Secondly, the LL problems do not include any binary decision variables since the derivation of the equivalent KKT optimality conditions is mathematically possible only when these problems are continuous and convex [77]. As a result, all relevant existing works [31, 34, 114, 115] make the unrealistic assumption that EV can adjust their charging / discharging power continuously between zero and maximum rates, and neglect that current EV battery and charger technologies allow discrete charging / discharging levels [117–119]. Therefore, employment of this approach may lead to sub-optimal pricing strategies for the EV aggregator.

To address these two fundamental limitations of previous works, this work resorts to *reinforcement learning* (RL) [80] which has recently emerged as an interesting alternative to MPEC formulations in electricity market modeling problems. In this modeling framework, the original bi-level optimization problem is not converted to a single-level, closed-form MPEC. Instead, it is solved in a recursive fashion; the EV aggregator gradually learns how to improve its pricing strategies (actions) by utilizing experiences acquired from its repeated interactions with the EV and the wholesale market. In other words, the aggregator does not rely on knowledge (perfect or imperfect) of the EV operating characteristics, but only on the observed EV responses (states of the environment). Furthermore, since this framework avoids the derivation of the equivalent KKT optimality conditions of the LL problems, it is capable of capturing the discrete charging / discharging levels of the EV.

However, previous works employing RL in electricity pricing design problems [39, 120, 121] have employed conventional Q-learning algorithms and its variants [80]. This type of algorithms relies on look-up tables to approximate the action-value function for each possible state-action pair and thus requires dicretization of both state and action spaces. Therefore, it suffers severely from the curse of dimensionality; as the number of considered discrete states and actions increases, the computational burden grows exponentially, soon rendering the problem intractable. If on the other hand a small number of discrete states and actions are considered, the feedback the aggregator receives regarding the impact of its pricing actions on the EV is distorted and the feasible action space is adversely affected, leading to sub-optimal pricing decisions. This challenge is aggravated in the setting of the examined problem, since both states of the environment (wholesale prices) and actions (retail prices) are not only continuous, but also multi-dimensional (due to the multi-period nature of the pricing problem).

In the context of addressing such dimensionality challenges, recent work in the emerging area of *deep reinforcement learning* (DRL) has proposed the deep Q network (DQN) method, which employs a deep neural network (DNN) to approximate the action-value function, and constitutes an extension of Q-learning to the multi-dimensional continuous state space [87]. Numerous recent papers have applied the DQN method to different electricity system problems [122–124], although its application in pricing design problems has yet to be explored. However, although this work has demonstrated high quality performance of the DQN method in problems with continuous state spaces, its performance in problems with continuous action spaces is less satisfactory because the employed DNN is trained to produce discrete action-value estimates rather than continuous actions [88], which significantly hinders its effectiveness in addressing the examined pricing design problem, since the aggregator's pricing actions are continuous and multi-dimensional.

As discussed above, Table 5.1 summarizes and compares the main characteristics of these works with respect to our work in this chapter.

Table 5.1 Summary of existing literature associated with the examined problem.

Paper	Optimization model	Require perfect knowledge	Ability to model discrete charging levels	State space	Action space
[31]	Bi-level	Yes	No	-	-
[34]	Bi-level	Yes	No	-	-
[114]	Bi-level	Yes	No	-	-
[115]	Bi-level	Yes	No	-	-
[120]	Q-learning	No	Yes	Discrete	Discrete
[121]	Q-learning	No	Yes	Discrete	Discrete
[39]	Q-learning	No	Yes	Discrete	Discrete
[122]	DQN	No	Yes	Continuous	Discrete
[123]	DQN	No	Yes	Continuous	Discrete
[124]	DQN	No	Yes	Continuous	Discrete
This work	PDDPG	No	Yes	Continuous	Continuous

5.3 Approach

This chapter proposes a new bi-level optimization problem for modeling the strategic EV pricing problem, which, in contrast with the existing literature, considers the V2G capability of EV and the discrete nature of their charging / discharging levels. However, as discussed in Section 5.2, the existing literature on EV pricing [31, 34, 114, 115] solves the formulated bi-level optimization problem by converting it to a single-level MPEC, through the replacement of the LL problems by their equivalent KKT optimality conditions. This reformulation is possible since the EV response optimization (LL) problems in this literature are convex as they assume that EV can adjust their charging / discharging power continuously. This solution approach is not applicable to our bi-level optimization problem since the aggregator does not acquire the perfect and imperfect knowledge of the EV, and the LL EV response problems are non-convex as they include the binary decision variables in order to capture the discrete nature of EV charging / discharging.

In order to address this challenge, we adopt a RL-based methodology, since it avoids the derivation of the equivalent KKT optimality conditions of the LL problem and solves the above bi-level optimization problem in a recursive fashion. Furthermore, in contrast with previous RL methods, a novel DRL method is developed to sets up the problem in multi-dimensional continuous state and action spaces. This method is named *prioritized deep deterministic policy gradient method* (PDDPG), as it is founded on the combination of deep deterministic policy gradient (DDPG) principles with a prioritized experience replay (PER) strategy.

5.3.1 Modeling Assumptions

For clarity reasons, the main assumptions behind the proposed model are outlined below:

1. The decision making problem of the examined aggregator considers the interaction with the number of flexible EV and inflexible EV (to which it sells energy or from which it buys energy when flexible EV exhibit V2G capability) with the overall aim of maximizing its profit.

2. It is assumed that the examined aggregator serves a relatively small population of EV and therefore its decisions do not affect the wholesale prices, which are thus

treated as fixed, exogenous parameters of the problem, i.e., the examined aggregator is assumed to be a price-taker in the wholesale market.

3. It is practical to assume that the aggregator has imperfect knowledge of the EV operation characteristics (especially their travelling patterns), when considering that the current penetration of EV is limited. Thus, the conventional approach of MPEC assuming that market players have knowledge of the computational algorithm of its examined problem (formulated as LL problem) generally cannot be applied here, and this motivates our work into the reinforcement learning approach.

4. In contrast with the traditional fixed pricing or time-of-use pricing regimes where the offered retail prices are flat throughout the examined daily horizon or during certain intervals of this horizon (e.g. peak and off-peak periods), the examined strategic aggregator can offer hour-specific retail prices to the served EV. In order to prevent the aggregator from making excessive profits at the expense of the EVs' cost, regulatory constraints are imposed on the maximum and average retail prices it can offer to its served EV [49].

5. The flexibility of served EV is represented as a discrete nature of the EV charging / discharging levels. In particular, most of the existing works consider the coordination of EV charging based on the assumption that EVs can adjust their charging power continuously between zero and their maximum charging rates (i.e., continuous charging). However, due to the limitations of the current battery technology (e.g., the lithium-ion battery) and EV charger technology (e.g., the constant-current constant voltage approach), EVs can only draw an approximately constant power during charging periods. Therefore, it is of practical importance to investigate the coordination of EV charging based on the discrete charging method.

6. Beyond the smart charging capability of EV (ability to flexibly select the periods when they buy energy for charging), their V2G capability is also considered (ability to sell energy through discharging).

7. The inflexible EV is assumed to start charging with the fixed charging rate immediately after it is connected to the grid until it covers its daily travelling energy requirements.

5.3.2 Structure of the Bi-level Optimization Model

In order to comprehensively capture the interactions between the aggregator, its flexible and inflexible EV, the proposed model is formulated as a bi-level optimization problem, the structure is illustrated in Fig. 5.1.

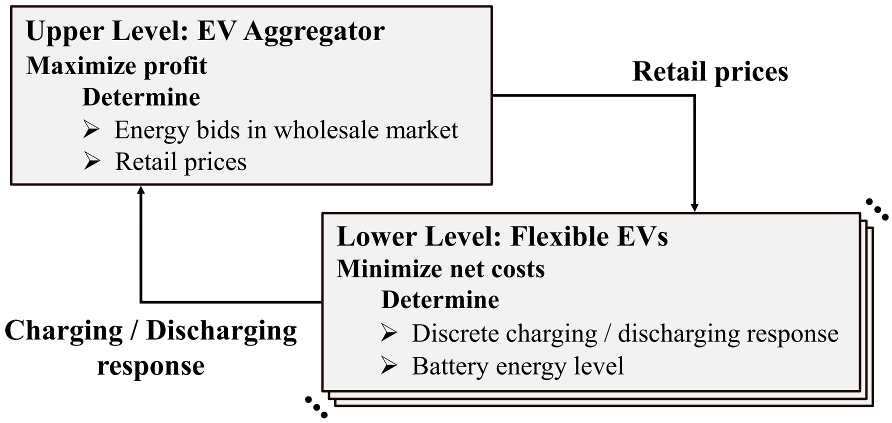

Fig. 5.1 Bi-level structure of the proposed pricing EV problem.

1. This bi-level model consists of an upper-level (UL) problem and multiple lower-level (LL) problems.
2. The examined aggregator behaves strategically through its strategic retail pricing and bidding decisions made at the UL problem. The objective of the UL problem is to maximize the net profit of the EV aggregator, which is defined as the difference between i) its net revenue from selling energy to its served inflexible and flexible EV, and ii) the net cost of buying energy from the wholesale market, and is subject to i) the maximum, minimum and average retail prices, and ii) the balance constraint between the energy sold / purchased to / from the EVs and the energy bought / sold from / to the wholesale market at each time period.
3. Each LL problem is considered representing the charging / discharging response of each flexible EV with the objective of minimizing its individual net cost, which is defined as the difference between i) its cost of buying energy from the EV aggregator for charging, and ii) its revenue from selling energy to the EV aggregator through discharging, and is subject to i) the EV discrete charging / discharging constraints, and ii) the energy constraints in the EV battery.
4. The UL problem and the multiple LL problems of Fig. 5.1 are all coupled, since the retail prices determined by the EV aggregator (UL problem) affect the response

of the flexible EVs (as they constitute part of the objective function of LL), while this response affects the retailer's decision making (as the EVs' demand constitutes part of the objective function and the energy balance constraint between the retail and wholesale markets of the UL).

5.4 Formulation

As discussed in Section 5.1 and following the existing literature on EV pricing [31, 34, 114, 115], the examined problem is formulated as a bi-level optimization problem. However, this bi-level problem is characterized by two important differences with respect to [31, 34, 114, 115]: a) beyond the smart charging capability of EV (ability to flexibly select the periods when they buy energy for charging), their V2G capability is also considered (ability to sell energy through discharging), and b) the charging / discharging levels of the EV are discrete and cannot be continuously adjusted.

5.4.1 Decision Making of Strategic EV Aggregator

The UL problem optimizes the pricing decisions of the examined aggregator, who acts as an intermediary entity between the EV and the wholesale market; this problem is formulated as follows:

$$\max_{\{\lambda_t^r, P_t^w\}} \sum\nolimits_{i,t} \lambda_t^r P_{i,t}^{inf} + \sum\nolimits_{j,t} \lambda_t^r C_{j,t} - \sum\nolimits_{j,t} \lambda_t^r D_{j,t} - \sum\nolimits_{t} \lambda_t^w P_t^w \tag{5.1}$$

subject to:

$$\lambda^{min} \leq \lambda_t^r \leq \lambda^{max}, \forall t \tag{5.2}$$

$$\sum\nolimits_t \lambda_t^r \,/\, |T| = \sum\nolimits_t \lambda_t^w \,/\, |T| \tag{5.3}$$

$$\sum\nolimits_i P_{i,t}^{inf} + \sum\nolimits_j C_{j,t} - \sum\nolimits_j D_{j,t} = P_t^w, \forall t \tag{5.4}$$

The objective function (5.1) maximizes the overall profit of the aggregator, which includes the following components: i) its revenue from selling energy to its inflexible EV, i.e. EV without smart charging or V2G capabilities (first term), ii) its revenue from selling energy to its flexible EV (second term), iii) its cost of buying energy from its flexible EV with V2G capability (third term), and iv) its net cost in the wholesale market, i.e. its cost / revenue of buying / selling energy from / to the wholesale market (fourth term). It is assumed that the examined aggregator serves a relatively small population of EV and therefore its decisions do not affect the wholesale prices λ_t^w, which are thus treated as fixed, exogenous parameters of the problem (i.e. the aggregator is assumed to be a price-taker in the wholesale market).

Following the formulation employed in [30], the maximum, minimum and average retail prices offered by the aggregator are subject to the regulatory constraints (5.2) and (5.3), which aim at preventing the aggregator from exploiting the EV and making excessive profits. Constraints (5.4) express the energy balance constraints of the aggregator; the net energy traded with its EV and the net energy traded with the wholesale market are equal at each period.

5.4.2 Charging / Discharging Response of Flexible EV

Each of LL problems optimizes the charging / discharging response of each flexible EV j to the retail prices offered by the aggregator, and is formulated as follows:

$$\min_{\{V_{j,t}^c, V_{j,t}^d, C_{j,t}, D_{j,t}, E_{j,t}\}} \sum_t \lambda_t^r C_{j,t} - \sum_t \lambda_t^r D_{j,t} \tag{5.5}$$

subject to:

$$V_{j,t}^c, V_{j,t}^d \in \{0,1\}, \forall t \tag{5.6}$$

$$V_{j,t}^c + V_{j,t}^d \leq 1, \forall t \tag{5.7}$$

$$C_{j,t} = V_{j,t}^c \, P_{j,t}^{max}, \forall t \tag{5.8}$$

$$D_{j,t} = V_{j,t}^d \, P_{j,t}^{max}, \forall t \tag{5.9}$$

$$P_{j,t}^{max} = \begin{cases} P_j^{max}, & \text{if } t \in T_j^{gr}; \\ 0, & \text{otherwise.} \end{cases} \tag{5.10}$$

$$E_{j,t} = E_{j,t-1} + C_{j,t} \, \eta_j^c \, \Delta t - D_{j,t}/(\eta_j^d \, \Delta t) - E_{j,t}^{tr}, \forall t \tag{5.11}$$

$$E_j^{min} \leq E_{j,t} \leq E_j^{max}, \forall t. \tag{5.12}$$

The objective function (5.5) minimizes the net cost of flexible EV j, which is defined as the difference between i) its cost of buying energy from the aggregator for charging (first term), and ii) its revenue from selling energy to the aggregator through discharging (second term).

In order to capture the discrete nature of EV charging / discharging, constraints (5.6)-(5.9) are employed. The binary variables $V_{j,t}^c$ and $V_{j,t}^d$ (5.6) indicate whether EV j

charges ($V^c_{j,t} = 1$), discharges ($V^d_{j,t} = 1$), or remains idle ($V^c_{j,t} = V^d_{j,t} = 0$) at period t, with constraint (5.7) ensuring that charging and discharging cannot happen simultaneously. Following the assumption of [117–119], if EV j charges or discharges at period t, then it does so based on a fixed power rate, equal to its maximum power rate at this period $P^{max}_{j,t}$ (5.8)-(5.9).

The latter parameter is defined by (5.10); it is either equal to the power rating of the battery if the EV j is connected to the grid at period t, or equal to zero if it is not (ensuring that charging or discharging cannot happen when the EV is not connected to the grid). Constraints (5.11) express the energy balance in the EV battery, including the energy required for travelling purposes as well as charging and discharging losses. Constraints (5.12) express the minimum and maximum limits of the battery's energy content.

5.4.3 RL Formulation

In this section, we detail the RL formulation of the examined EV aggregator pricing problem, the key elements of which are outlined in Fig. 5.2:

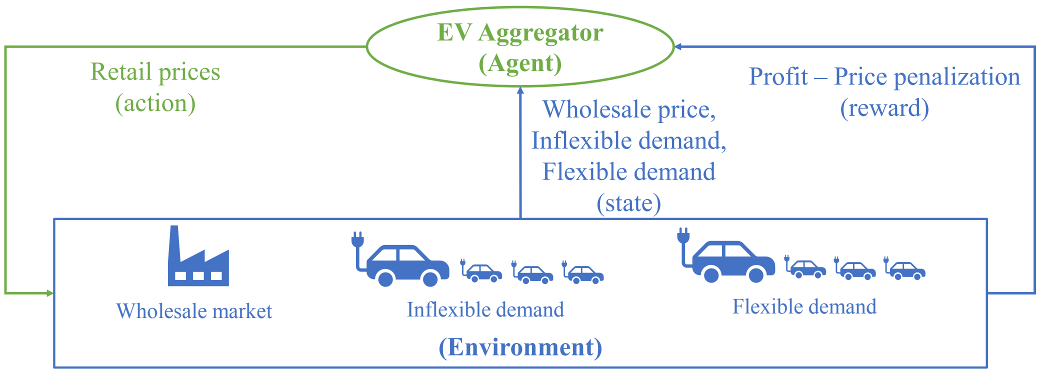

Fig. 5.2 Agent-environment interactions in the examined problem.

1) ***Agent***: The examined EV aggregator constitutes the agent, which gradually learns how to improve its retail pricing decisions by utilizing experiences from its repeated interactions with the environment.

2) ***Environment***: The environment consists of the EV (both inflexible and flexible ones) and the wholesale market, with both of which the EV aggregator interacts.

3) ***State***: The state vector s plays the role of a feedback signal for the agent regarding the impact of its action on the environment. In this problem, this is a $3 \times |T|$-dimensional vector which consists of the wholesale market prices, the demand of

inflexible EV and the net demand of flexible EV, i.e. $s = [\lambda^w_{1:|T|}, \sum_i P^{inf}_{i,1:|T|}, \sum_j (C_{j,1:|T|} - D_{j,1:|T|})] \in \mathcal{S}$.

4) ***Action***: The action is a $|T|$-dimensional continuous vector which includes the retail prices offered by the aggregator, i.e. $a = [\lambda^r_{1:|T|}] \in \mathcal{A}$. These retail prices should satisfy the regulatory constraints (5.2) and (5.3). However, a fundamental limitation of RL is that constraints coupling action dimensions (i.e. the time-coupling average retail price constraint (5.3) in the examined problem) cannot be directly satisfied [125]. This limitation is addressed as explained in the following point.

5) ***Reward***: From a physical perspective, the reward of the aggregator should be set equal to its overall profit (5.1). However, in order to address the limitation discussed in the previous point, we add a penalty term in the reward (5.13), penalizing the extent of violation of constraint (5.3), i.e. the absolute difference between the average retail price and the average wholesale price, with W denoting the (positive) penalty weighting constant. By employing a large enough value for W, the RL algorithm is incentivized to select actions that satisfy constraint (5.3).

$$r = (5.1) - W * \Big| \sum_t \lambda^r_t / |T| - \sum_t \lambda^w_t / |T| \Big|. \tag{5.13}$$

5.4.4 State-of-the-art RL Methods

Before detailing the proposed RL method, we briefly discuss the two state-of-the-art RL methods employed in the electricity system literature (Section 5.2).

Q-Learning: The Q-value function can be described in a recursive format according to the *Bellman equation* [126]:

$$Q_\pi(s_\tau, a_\tau) = \mathbb{E}[r_\tau + \gamma\, Q_\pi(s_{\tau+1}, \pi(s_{\tau+1}))] \tag{5.14}$$

The Bellman equation indicates that the Q-value can be updated by *bootstrapping*, i.e. we can improve the future estimate of Q by using the current estimate of Q. This serves as the foundation of Q-learning [83], a form of *temporal difference* (TD) learning [80], where the Q-value is updated after taking action a_τ at state s_τ and observing reward r_τ and resulting state $s_{\tau+1}$, a detailed updating procedure of the Q-Learning method is outlined in Algorithm 1.

$$Q(s_\tau, a_\tau) \;\leftarrow\; Q(s_\tau, a_\tau) + \alpha\, \delta_\tau \tag{5.15}$$

$$\delta_\tau = r_\tau + \gamma \ \max_{a_{\tau+1}} Q(s_{\tau+1}, a_{\tau+1}) - Q(s_\tau, a_\tau) \tag{5.16}$$

where $\alpha \in [0, 1]$ is the step size, δ_τ is the correction of the estimation of the Q-value function (known as the TD error), and $r_\tau + \gamma \max_{a_{\tau+1}} Q(s_{\tau+1}, a_{\tau+1})$ expresses the target Q-value at time step τ.

The final Q-Learning method is outlined in Algorithm 1.

Algorithm 1 Updating procedure of Q-Learning

```
1: Initialize the Q-table Q(s, a) for s ∈ S and a ∈ A.
2: for episode = 1 : M^train do
3:    Selects random retail prices in the action space and the resulting net demand
      of flexible EV (solved in LL problems of Fig. 5.1), along with the demand
      of inflexible EV and the wholesale prices are used as the initial state of the
      environment for the current episode.
4:    for time step = 1 : T do
5:       Selects action a_t based on the ϵ-greedy policy [80]:
           with probability ϵ select a random action
           otherwise select a = argmax_a Q(s, a).
6:       Execute action a_τ in the environment, observe reward r_τ, and transit to the
         new state s_{τ+1}.
7:       Update the Q-table using (5.15) and (5.16).
8:    end for
9: end for
```

If the Q-value for each admissible state-action pair is visited infinitely often, and the learning rate α decreases over the time step τ in a suitable way, then as $\tau \to \infty$, $Q(s, a)$ converges with probability one to the optimal $Q_*(s, a)$ for all admissible state-action pairs [83].

Although Q-learning merits simplicity and its convergence is theoretically guaranteed, it suffers severely from the curse of dimensionality [80]. This is because Q-learning stores the Q-value function in a look-up table, which then necessitates the RL problem being set up in discrete state and action spaces. Therefore, as the number of considered discrete states and actions increases, the computational burden grows exponentially, soon rendering the problem intractable. For example, if we discretize each state and action dimension in 10 integer values, then for a 10-dimensional state and action vector, this leads to 10^{10} rows and 10^{10} columns of the look-up table. In other words, the granularity of the discretization of the state and action spaces affects significantly the performance of Q-learning. Specifically, a lower granularity results in poor generalization capabilities of Q-learning, whereas a higher granularity, despite helping Q-learning generalize better, leads to an exponential growth of the number of states and actions, and consequently impractical memory requirements.

DQN: In order to address the dimensionality challenges of tabular-based Q-learning, the DQN method [86, 87] employs a DNN, parameterized by θ, as a function approximator to represent the Q-value function in multi-dimensional continuous state space:

$$Q(s_\tau, a_\tau) \approx Q(s_\tau, a_\tau|\theta) \tag{5.17}$$

DQN takes as input a continuous state s_τ and outputs an estimate of the Q-value function for each discrete action and, when acting, selects the maximally valued output at a given state. Prior to DQN, the use of large, non-linear function approximators (such as DNNs) for learning the Q-value function has generally been averted since theoretical convergence guarantees are impossible, and the learning tends to be unstable [80]. However, such non-linear approximators prove to be essential for the agent to learn and generalize on a multi-dimensional continuous state space [88]. Nevertheless, DQN is able to learn the Q-value function using DNN in a stable and robust fashion. This is enabled by two innovations: the use of the experience replay $\mathcal{R}$ [86, 87] and the target network $Q_{\theta'}(s, a)$ [87]. In brief, the former stores gathered experiences (an experience is a transition tuple $(s_\tau, a_\tau, r_\tau, s_{\tau+1})$) in a replay buffer $\mathcal{R}$ and randomly samples a minibatch (of size N) of experiences to to train the DNN, enabling decorrelation of consecutively generated training samples. The later temporarily freezes the Q-value target during training, thereby stabilizing the learning process. A more detailed discussion of these two mechanisms is presented in Section 5.4.5.

The training procedure of the DQN method is outlined in Algorithm 2. At each time step t, we sample a minibatch of N training experiences $\{(s_n, a_n, r_n, s_{n+1})\}_{n=1}^N$ from $\mathcal{R}$, the training of the DNN is based on temporal difference (TD) learning through the minimization of the following loss function, representing the mean-squared TD error:

$$L_\theta = \frac{1}{N}\sum_{n=1}^{N}\Big(r_n + \gamma \max_{a_{n+1}} Q_{\theta'}(s_{n+1}, a_{n+1}) - Q_\theta(s_n, a_n)\Big)^2 \tag{5.18}$$

and the following update is applied to the DNN weights θ, where α is the learning rate for the gradient descent algorithm. The target network weights θ' is updated with the online network weights θ every C time steps to stabilize learning.

$$\theta \leftarrow \theta + \alpha \nabla_\theta L_\theta. \tag{5.19}$$

The final DQN method is outlined in Algorithm 2.

Algorithm 2 Training procedure of DQN

```
Initialize the online Q network with random weights θ.
Initialize the target Q network with weights θ′ ← θ.
for episode = 1 : M do
    Selects random retail prices in the action space and the resulting net demand
    of flexible EV (solved in LL problems of Fig. 5.1), along with the demand
    of inflexible EV and the wholesale prices are used as the initial state of the
    environment for the current episode.
    for time step = 1 : T do
        Selects action a_τ based on the ϵ-greedy policy [80]:
          with probability ϵ select a random action
          otherwise select a = argmax_a Q_θ(s, a).
        Execute action a_τ in the environment, observe reward r_τ using (5.13), and
        transit to the new state s_{τ+1}.
        Store experience (s_τ, a_τ, r_τ, s_{τ+1}) in R.
        Sample uniformly a minibatch of transitions {(s_n, a_n, r_n, s_{n+1})}_{n=1}^N from R.
        Update the online network using (5.18) and (5.19).
        Update the target network with θ′ with θ′ ← θ every C time steps.
    end for
end for
```

Although DQN performs well in problems with continuous state space, the curse of dimensionality persists. This is owing to the fact that the DNN are trained to output discrete estimates of the Q-values rather than continuous actions. Therefore, to address problems with a continuous action domain, dicretization of the action space is required. Furthermore, even if a good approximation of the Q-value function can be obtained, it is still challenging to find a continuous action that corresponds to the highest Q-value at a given state. This motivates investigation of more sophisticated RL methods which facilitate effective learning in multi-dimensional continuous state and action spaces, as prescribed by the examined pricing problem.

5.4.5 Proposed PDDPG Method

In order to address the limitations of the DQN method in problems with continuous action spaces (Section 5.2), we propose the novel PDDPG method, the overall workflow of which is illustrated in Fig. 5.3.

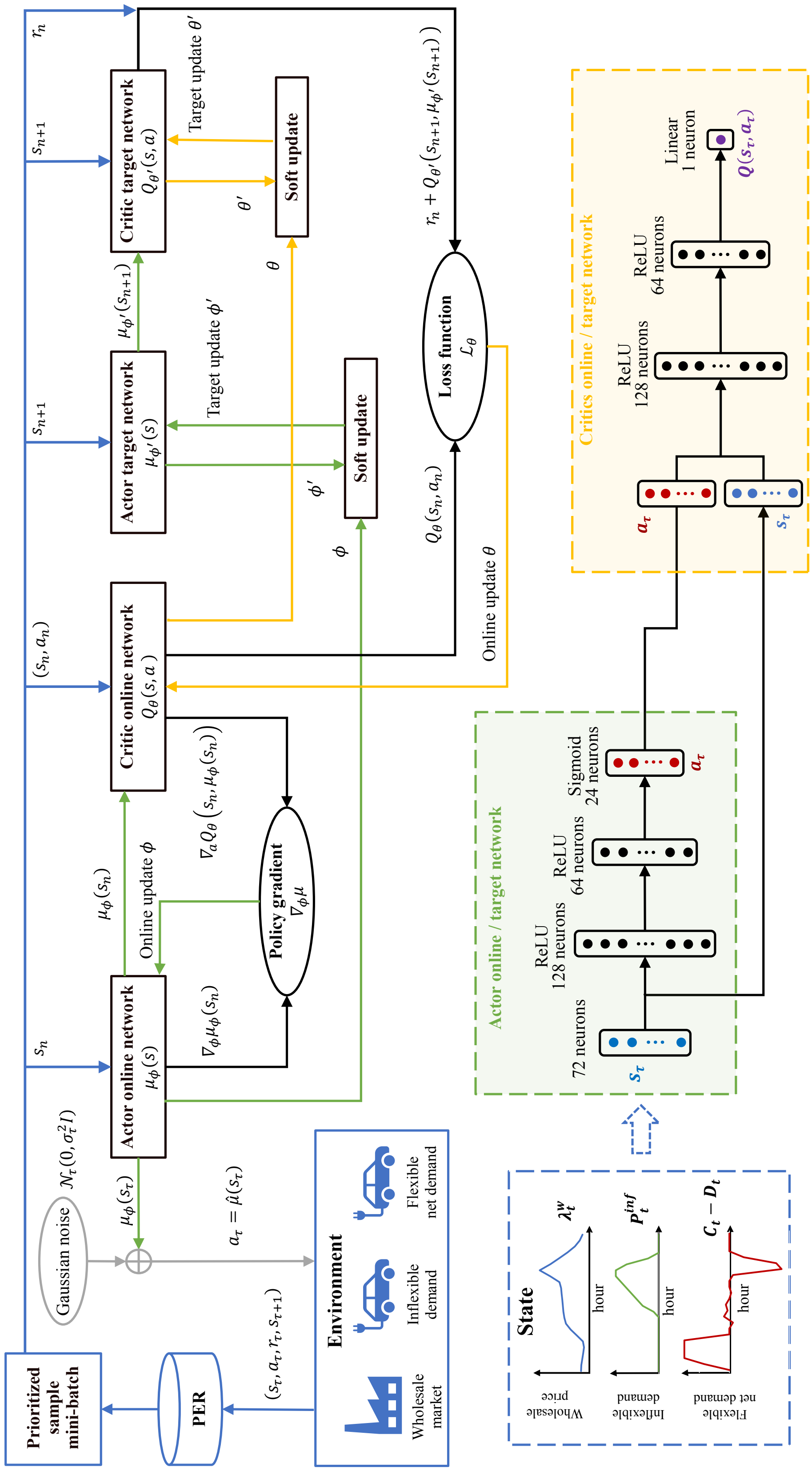

Fig. 5.3 Workflow of proposed PDDPG method.

PDDPG features an *actor-critic* architecture and employs two DNNs for different purposes [88]. The *critic network* Q_θ, parameterized by θ, takes as input a state s_τ and action a_τ and outputs a scalar estimate of the Q-value function $Q_\theta(s_\tau, a_\tau)$. The *actor network* μ_ϕ, parameterized by ϕ, takes as input a state s_τ and implements the policy improvement task, updating the policy with respect to the estimated Q-value function and outputting a continuous action $\mu_\phi(s_\tau)$.

Regarding the policy improvement, the common approach adopted in Q-learning and DQN involves a greedy maximization of the Q-value function, i.e., $\mu(s_\tau) = \operatorname{argmax}_{a_\tau} Q(s_\tau, a_\tau)$ (Section 5.4.4). However, greedy policy improvement tends to be intractable in multi-dimensional continuous action spaces as it requires the global maximization of the Q-value function at every time step. In this context, the proposed method employs the actor network μ to generate an action $\mu_\phi(s_\tau)$ for the current state. The critic network then implements the policy evaluation task, appraising the policy by producing an estimate of the Q-value function with TD learning. This is achieved through the minimization of the following loss function:

$$\mathcal{L}(\theta) = \delta_\tau^2 \tag{5.20}$$

$$\delta_\tau = r_\tau + \gamma\, Q_\theta(s_{\tau+1}, \mu(s_{\tau+1})) - Q_\theta(s_\tau, a_\tau). \tag{5.21}$$

where δ_τ and $r_\tau + \gamma\, Q_\theta(s_{\tau+1}, \mu(s_{\tau+1}))$ express the TD error and the target Q-value at time step τ, respectively. Instead of globally maximizing $Q_\theta(s_\tau, a_\tau)$, the critic calculates gradients $\nabla_a Q_\theta(s_\tau, a_\tau)$ which indicate directions of change of action resulting in higher estimated Q-values. These gradients are calculated via back-propagation through the critic, which is more efficient in computational terms than solving an optimization problem in continuous action space. In order to update the actor μ_ϕ, these gradients are placed at the actor's output layer and then are back-propagated through the network. For a given state s_τ, a forward pass through the actor produces an action a_τ that the critic appraises, and the resulting policy gradients $\nabla_\phi\, \mu$ are employed to update the actor:

$$\nabla_\phi \mu = \nabla_a Q_\theta(s, a)|_{s=s_\tau, a=\mu(s_\tau)} \nabla_\phi \mu_\phi(s)|_{s=s_\tau}. \tag{5.22}$$

In RL, maintaining an appropriate balance between *exploration* and *exploitation* is deemed imperative in selecting the agent's actions. Through exploration, the agent collects more information by trying out different actions in the action space. Through exploitation, the agent learns to make the best decision given the available information. A significant advantage of the proposed method lies in decoupling the exploration problem from the learning algorithm. Specifically, we construct an exploration policy $\hat{\mu}(s_\tau)$ by adding a random Gaussian noise $\mathcal{N}_\tau(0, \sigma^2 I)$ to the actor's output $\mu_\phi(s_\tau)$ in order to aid the agent in exploring the environment thoroughly:

$$\hat{\mu}(s_\tau) = \mu_\phi(s_\tau) + \mathcal{N}_\tau(0, \sigma^2 I). \tag{5.23}$$

RL tends to exhibit unstable learning behavior when used in combination with a DNN as the Q-value function approximator [87]. First of all, since the online network $Q_\theta(s_\tau, a_\tau)$ being updated is also employed in evaluating the target value $r_\tau + \gamma\, Q_\theta(s_{\tau+1}, \mu(s_{\tau+1}))$ (as shown in (5.21)), the Q-value update is prone to oscillations. An effective approach to tackle this instability is to introduce a target network[87] for both actor and critic networks, denoted as $\mu_{\phi'}(s_\tau)$ and $Q_{\theta'}(s_\tau, a_\tau)$, respectively, and employ them to evaluate the target values. The weights of these target networks (ϕ' and θ') are updated by slowly tracking the weights of online networks (ϕ and θ) as $\phi' \leftarrow \nu\phi + (1-\nu)\phi'$ and $\theta' \leftarrow \nu\theta + (1-\nu)\theta'$ with $\nu \ll 1$. The logic behind this *soft update* lies in restricting the target values (for both actor and critic) to change slowly so as to stabilize the learning process.

During the learning process, experiences are generated as the agent sequentially interacts with the environment, meaning that these samples are temporally correlated, which hinders the application of deep learning algorithms. An effective solution to remedy this issue is to incorporate an experience replay buffer which is a pool that stores past experiences and uniformly samples a minibatch to update the actor and critic at each time step. Mixing recent with previous experiences contributes to diminishing the temporal correlations presented in the replayed experiences. Additionally, the experience replay allows samples to be reused, and thereby increases the sampling efficiency [87].

To further enhance the sampling efficiency of the original experience replay mechanism and thereby accelerate the learning process, we propose the use of the *prioritized experience replay* method [127]. In this method, the magnitude of TD-error is used as a measure of the correction for the Q-value estimation. In other words, it suggests to what extent an agent could learn from an experience. Experiences with large positive TD-errors are associated with very successful attempts, while experiences with large negative TD-errors signal the conditions where the agent's actions are highly erroneous and therefore that the states of these conditions are inadequately learned. Prioritizing replaying of these experiences during training enables the agent to improve its policy more quickly from very successful attempts and also prevents the agent from selecting the unfavorable actions in some states, and thereby improves the quality of the policy learned.

Consider a prioritized experience replay buffer $\mathcal{PR}$ of size N_{PR}, we define the probability P_n of sampling experience n based on the absolute TD error as:

$$P_n = p_n^{\beta_1} / \sum_m p_m^{\beta_1} \tag{5.24}$$

where β_1 controls the extent of prioritization used, $p_n = 1/rank_n$ is the priority assigned to experience n, and $rank_n$ indicates the rank of experience n when the $\mathcal{PR}$ is ordered according to the absolute TD error $|\delta_n|$, which is defined as:

$$|\delta_n| = \left| r_n + \gamma\, Q_{\theta'} \Big(s_{n+1}, \mu_{\phi'}(s_{n+1}) \Big) - Q_\theta(s_n, a_n) \right| \tag{5.25}$$

However, since experiences with high $|\delta|$ are more regularly replayed, this practice alters the visitation frequency of some states and thus introduces bias. To correct this bias, we resort to importance-sampling (IS) weights:

$$W_n = (N_{PR} P_n)^{-\beta_2} / \max_m W_m \tag{5.26}$$

where β_2 controls the extent of correction used. These weights are then incorporated in the computation of the critic loss by substituting δ_n with $W_n \delta_n$. For stability reasons, the IS weights are normalized by $\max_m W_m$.

Finally, by employing the target network and the prioritized experience replay, the critic loss (5.20) can be restated as the weighted mean-squared TD error calculated based on the training data, i.e. a minibatch of prioritized sampled N experiences with priority p_n and IS weights W_n.

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{n=1}^{N} W_n \delta_n^2 \tag{5.27}$$

The policy gradient for the actor update (5.22) can be restated in a similar fashion as:

$$\nabla_\phi \mu = \frac{1}{N} \sum_{n=1}^{N} \nabla_a Q_\theta(s_n, a)|_{a=\mu_\phi(s_n)} \nabla_\phi \mu_\phi(s_n) \tag{5.28}$$

The following updates are applied to the weights of the online critic and actor networks, respectively, where α^θ and α^ϕ are the learning rates of the gradient decent algorithm:

$$\theta \leftarrow \theta + \alpha^{\theta} \nabla_{\theta} L_{\theta} \tag{5.29}$$

$$\phi \leftarrow \phi + \alpha^{\phi} \nabla_{\phi} J(\mu_{\phi}) \tag{5.30}$$

The target critic and actor networks are updated to gradually track the online critic and actor networks according to:

$$\theta' \leftarrow \nu\theta + (1 - \nu)\theta' \tag{5.31}$$

$$\phi' \leftarrow \nu\phi + (1 - \nu)\phi' \tag{5.32}$$

The final PDDPG method is outlined in Algorithm 3.

Algorithm 3 Proposed PDDPG Algorithm

Initialize the online critic and actor networks with random weights θ and ϕ, respectively.
Initialize the target critic and actor networks with weights $\theta' \leftarrow \theta$ and $\phi' \leftarrow \phi$, respectively.
for episode $= 1 : M$ **do**
    Select random retail prices in the action space; the resulting net demand of flexible EV (solved in LL problems of Fig. 5.1), along with the demand of inflexible EV and the wholesale prices, are used as the initial state of the environment for the current episode.
    Initialize a random Gaussian exploration noise $\mathcal{N}_{\tau}$.
    for time step $= 1 : T$ **do**
        Select retail pricing decisions using (5.23).
        The new state $s_{\tau+1}$ is determined and the reward r_{τ} is calculated through (5.13).
        Store, in $\mathcal{PR}$, experience $(s_{\tau}, a_{\tau}, r_{\tau}, s_{\tau+1})$ and set $p_{\tau} = \max_{n<\tau} p_n$.
        for $n = 1 : N$ **do**
            Sample experience n with probability P_n in (5.24).
            Compute the absolute TD-error $|\delta_n|$ using (5.25).
            Compute the IS weights W_n using (5.26).
            Update the priority p_n according to $|\delta_n|$.
        end for
        Update the critic network using (5.27) and (5.29).
        Update the actor network using (5.28) and (5.30).
        Update the target networks using (5.31) and (5.32).
    end for
end for

5.5 Case Studies

5.5.1 Test Data and Implementation

The following case studies examine the EV aggregator pricing problem (Section 3.4) in the context of a single day with hourly resolution, i.e. $T = \{1, 2, ..., 24\}$. The examined aggregator serves a population of 1,000 EV, divided in 25 different types (of 40 EV each) according to their travelling times, grid connection times and travelling energy requirements in the examined day. The travelling data is based on UK national surveys [128] while the EV are assumed connected to the grid between the end of their last trip and the start of their first trip of the day, in line with the home-charging paradigm. For space limitation reasons, this data is provided in the Appendix A.

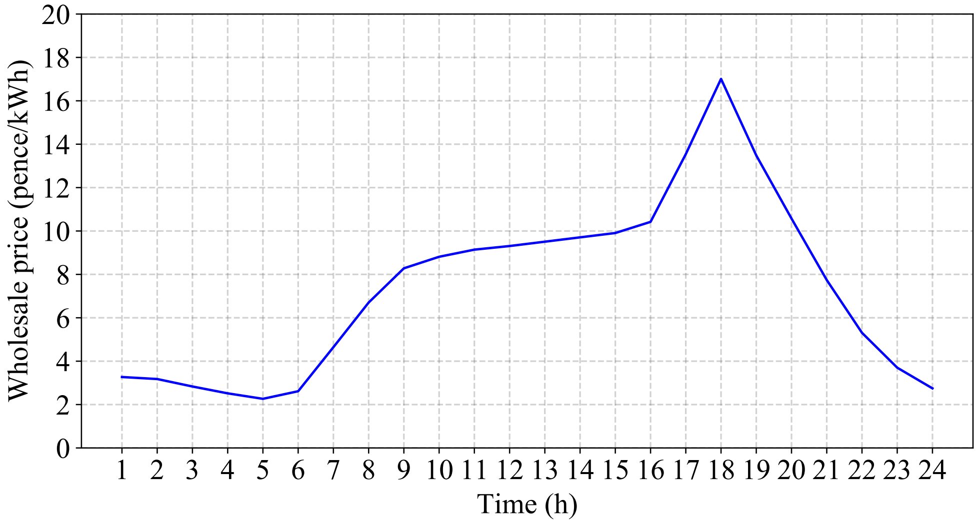

Fig. 5.4 Wholesale market price of the examined day.

The assumed values of the remaining technical parameters of the EV are P_j^{max} = 3 kW, E_j^{min} = 1.5 kWh, E_j^{max} = 15 kWh, $\eta_j^c = \eta_j^d$ = 0.93 for every flexible EV j. Each inflexible EV i is assumed to start charging (with the same, fixed rate of 3 kW) immediately after it is connected to the grid until it covers its daily travelling energy requirements; the charging power profile of the inflexible EV of each of the 25 types is also provided in the Appendix A. In the following studies, different scenarios are examined regarding the percentage of EV being flexible as well as the nature of their flexibility (smart charging capability only or combined smart charging and V2G capabilities). The assumed wholesale market prices across the examined day follow the pattern of a typical winter day in the UK, which are illustrated in Fig. 5.4. Finally, the regulatory minimum and maximum retail prices are assumed $\lambda^{min} = 2$ pence/kWh and $\lambda^{max} = 20$ pence/kWh, respectively.

The main implementation details of the 3 examined RL methods are discussed as:

Q-Learning: The examined problem needs to be set up in discrete state and action spaces. We discretize the hourly net demands of flexible EV (states) in 20 integer values and the hourly retail prices (actions) in 19 integer values (between $\lambda^{min} = 2$ pence/kWh and $\lambda^{max} = 20$ pence/kWh, with a step of 1 penny/kWh). Therefore, the aggregator employs 24 look-up tables, each of size 20×19 to store and update the Q-values of state-action pairs at each hour. Note that it is impractical to use a single look-up table of size $20^{24} \times 19^{24}$ to store the Q-values associated with different daily state-action pairs under the assumed discretization.

DQN: This method employs a DNN as an approximator that provides the Q-value estimate for each discrete action and, when acting, selects the action corresponding to the highest Q-value at a given state. Similar to the approach adopted in [123], in the examined problem the state is represented as a time-window of two adjacent hours, i.e., hour identifier t, and wholesale price, demand of inflexible EV, and net demand of flexible EV at hours $t-1$ and t, resulting in 7 neurons in the input layer of the DNN. The continuous action space is discretized in the same fashion as in Q-learning, resulting in 19 neurons in the output layer of the DNN.

PDDPG: We employ two DNNs i.e. the online and the target network, for both actor and critic (Fig. 5.3). The Adam optimizer is employed for learning the neural network weights with a learning rate $\alpha^{\mu} = 10^{-4}$ and $\alpha^{Q} = 10^{-3}$ for the actor and critic, respectively. For the critic, we use a discount factor of $\gamma = 0.7$, and include a L_2 regularization term in its loss function with a weight decay rate of 10^{-2} in order to avoid very large weights. We use $\nu = 0.001$ as the target network updating rate. The structure of the actor and critic networks is illustrated in Fig. 5.3. Both the actor and critic employ the rectified non-linearity (ReLU) for all hidden layers. The output layer of the actor is a sigmoid layer to bound the continuous actions. We train with a minibatch size of $N = 128$ and for $M = 1000$ episodes, with 20 days per episode. The parameters employed in the prioritized experience replay are set as $\beta = 0.6$ and $N_B = 10^4$. We use $W = 10^3$ for the penalty weighting constant in the reward (5.13).

The examined RL methods have been implemented using Tensorflow [129] in Python. The LL problem of Fig. 5.1 has been implemented using the Xpress Optimizer Python interface [130]. The case studies have been carried out on a computer with a 6-core 3.50 GHz Intel(R) Xeon(R) E5-1650 v3 processor and 32 GB of RAM.

5.5.2 Performance Comparison of RL Methods

The aim of the first set of studies lies in comparing the performance of the proposed PDDPG method against the state-of-the-art Q-Learning and DQN methods in the examined problem, considering a scenario where 25% of the EV of each type are assumed flexible with smart charging capability only. We generate 10 different random seeds, and for each one we train each of the three RL methods for 1000 episodes, with one episode consisting of 20 time steps (Algorithm 3). It should be mentioned that the random seed is used to fix the stochasticity in RL simulation (e.g., initial state of per episode, weights initialization of the DNNs), thus we generate 10 random seeds for the experiment to evaluate the robustness and variation of the proposed methods to the examined problem. For the detailed analysis, please refer to the standard deviation in Fig. 5.5 and Table 5.2.

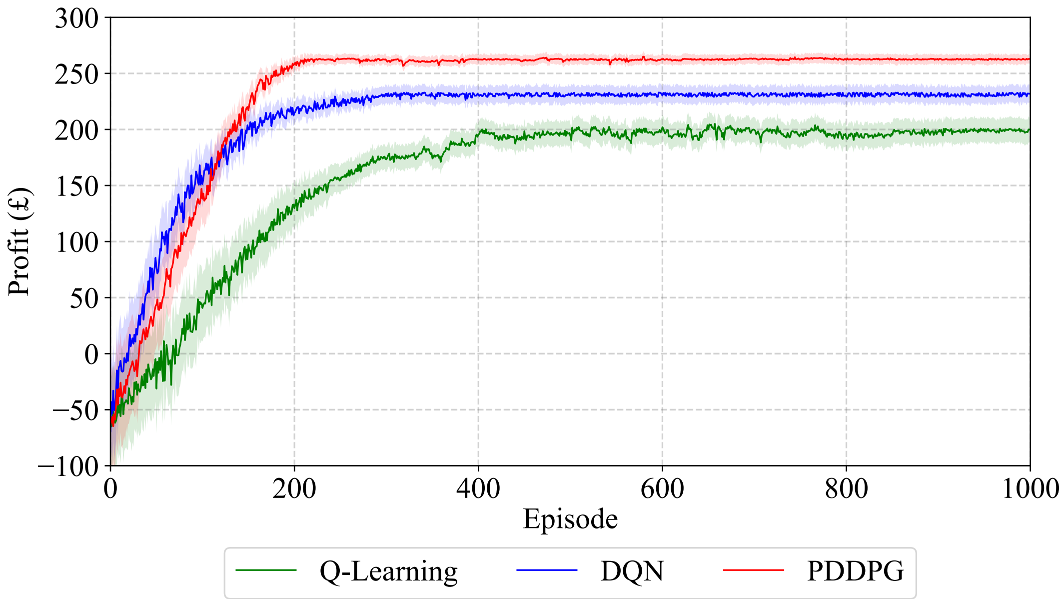

Fig. 5.5 Episodic average profit over 10 different random seeds for the examined RL methods.

Fig. 5.5 and Table 5.2 present the episodic average profit of the aggregator over the 10 different random seeds for each RL method. The mean μ and the standard deviation σ of the average profit over the 10 seeds at different episodes are illustrated in Fig. 5.5 through the solid lines and the shaded areas, respectively, and are also provided in Table 5.2. As demonstrated in Fig. 5.5, the average profit is negative during the initial learning stages, since the aggregator is gathering more experiences by randomly exploring different, not necessarily profitable, actions. However, as the learning process progresses and more experiences are collected, the average profit turns positive, it keeps increasing, and eventually converges for all three methods. This convergence is reflected in the stabilized mean and standard deviation of the average profit during the last learning stages (Fig. 5.5 and Table 5.2).

Table 5.2 Mean μ (£) and standard deviation σ (£) of average profit over 10 different random seeds at different episodes for examined RL methods.

		Episode					
		0	100	200	300	600	1000
Q-Learning	μ	-61.15	45.00	133.06	173.46	198.63	199.94
	σ	44.72	33.19	19.75	11.29	11.41	11.54
DQN	μ	-69.42	158.83	212.49	231.70	231.57	231.84
	σ	43.94	20.09	9.60	8.56	8.09	8.05
PDDPG	μ	-63.86	147.18	257.24	262.55	262.74	**262.89**
	σ	46.37	24.90	7.69	5.19	4.29	**4.14**

The proposed PDDPG method exhibits a larger standard deviation compared to Q-learning and DQN, and a slower learning pace compared to DQN during the initial learning stages. This effect is driven by the fact that the exploration of suitable pricing decisions in multi-dimensional continuous action space (performed by PDDPG) is more challenging than in discrete space (performed by Q-Learning and DQN). However, as the learning process progresses, the PDDPG method significantly outperforms the two state-of-the-art methods, and converges to the highest average profit (£262.89) and the lowest standard deviation (£4.14). In relative terms, PDDPG achieves 31% / 13% higher average profit and 64% / 49% lower standard deviation with respect to Q-Learning / DQN.

5.5.3 Impact of PER on Learning Performance

This section lies in demonstrating the impact of PER, by comparing the performance of the PDDPG method against the original DDPG method adopting uniform sampling (Section 5.4.5). Fig. 5.6 presents the performance of the examined methods in terms of their learning speed and policy quality. The solid lines illustrate the episodic average profit over 10 different random seeds for the two examined methods while the dashed lines indicate the number of episodes required to reach the baseline profit level (set as the episodic average profit of the DQN method at convergence).

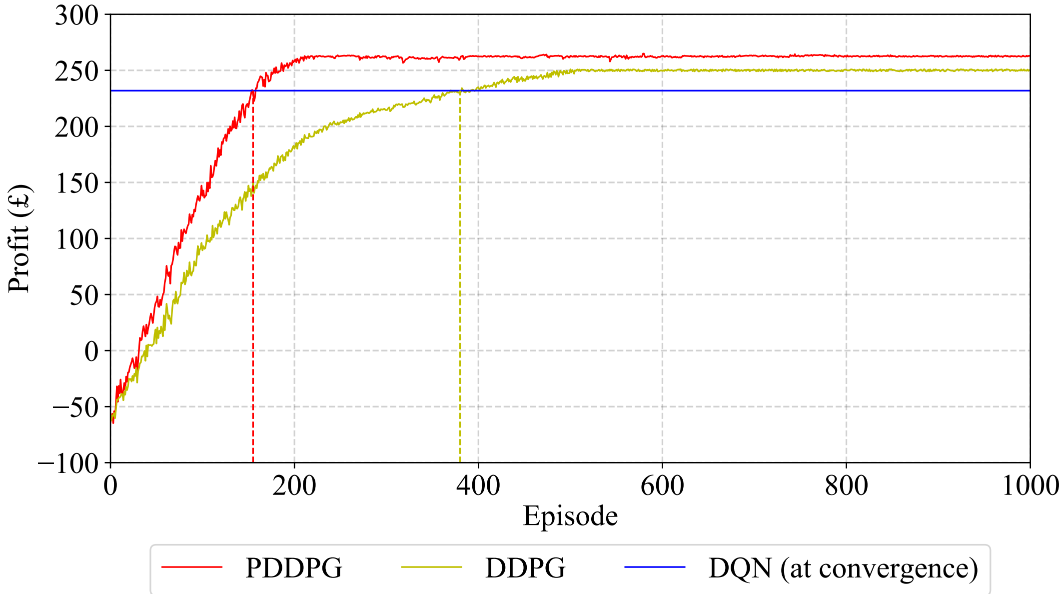

Fig. 5.6 Episodic average profit over 10 different random seeds for the DDPG and PDDPG methods..

As shown in Fig. 5.6, PDDPG and DDPG reach the baseline reward in approximately 155 and 380 episodes, respectively, suggesting that the learning speed of PDDPG is approximately 2.45 times as fast as DDPG. Furthermore, PDDPG results in more profitable pricing decisions, reflected by the approximately 5.08% higher average profit at convergence. The rationale behind this favourable learning speed and policy quality lies in the fact that PER more frequently replays experiences corresponding to higher TD-error which promises significant benefits in improving the agent's policy. On the other hand, such experiences may be rarely (or never) replayed when uniform sampling is adopted.

5.5.4 Continuous Action Vs. Discrete Action

The superior performance of the proposed method with respect to the most advanced state-of-the-art RL method (DQN) is driven by its ability to capture a continuous action space in contrast with the naive discretisation approach adopted by DQN. In order to investigate the physical significance of this methodological advantage of PDDPG, Figs. 5.7-5.8 and Table 5.3 compare the retail pricing decisions of the aggregator and the resulting demand of all EV at the convergence state of the DQN and PDDPG methods. As demonstrated in Figs. 5.7-5.8, the fundamental difference between the solutions obtained by the two methods lies in the retail pricing decisions at hours 1-13. From an economic perspective, the aggregator aims at attracting the demand of flexible EV (with smart charging capability only in the scenario examined in this Section) at the hours with the lowest wholesale market prices (hours 4 and 5) in order to reduce its cost in the wholesale market (fourth term of its profit (5.1)); therefore, the aggregator is motivated to offer lower retail prices at these two hours with respect to other hours of the day.

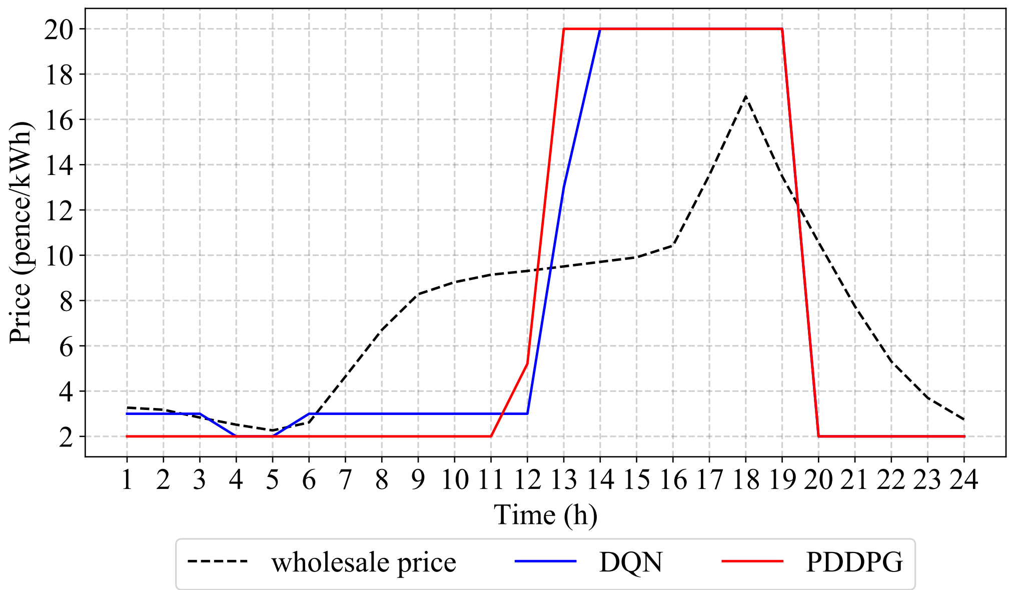

Fig. 5.7 Retail prices for DQN and PDDPG methods.

Table 5.3 Retail prices (pence/kWh) for DQN and PDDPG methods.

	Hour 3	**Hour 4**	**Hour 5**	**Hour 6**	**Hour 12**	**Hour 13**
λ_t^w	2.8329	**2.5171**	**2.2644**	2.6147	9.3071	9.5071
λ_t^r (DQN)	3	**2**	**2**	3	3	13
λ_t^r (PDDPG)	2.0006	**2.0003**	**2.0003**	2.0004	5.2109	19.9981

Under PDDPG, given that the pricing actions can continuously vary, the difference between the prices at these two hours and other, adjacent off-peak hours is marginally low (Table 5.3). Under DQN, given that the pricing actions can only vary in discrete steps (of 1 penny/kWh in the examined studies, Section 5.5.1), the aggregator is forced

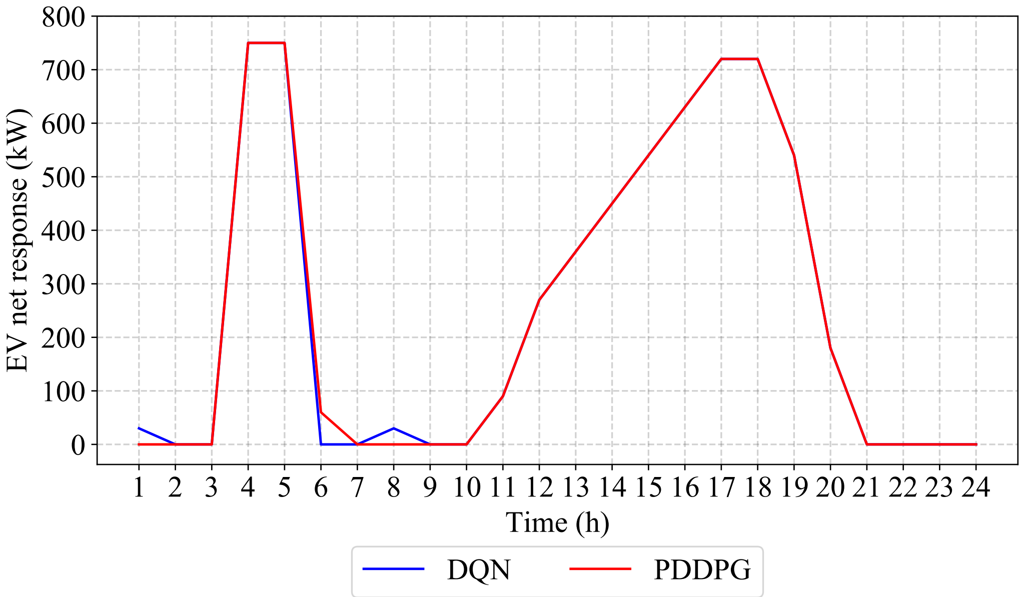

Fig. 5.8 EV net demand for DQN and PDDPG methods.

to differentiate the prices at these two hours with respect to adjacent off-peak hours by 1 penny/kWh; therefore, it offers the lowest allowable price $\lambda^{min} = 2$ pence/kWh at hours 4 and 5 and a price of 3 pence/kWh at the off-peak hours 1-3 and 6-11 (Fig. 5.7). Because of this substantial increase of the offered prices at hours 1-3 and 6-11 (with respect to PDDPG, where the respective prices are still very close to the lowest allowable level of 2 pence/kWh), and in order to satisfy the average retail price constraint (5.3), the aggregator is forced to reduce the offered prices at other, peak hours of the day (specifically, hours 12 and 13 as demonstrated in Fig. 5.7 and Table 5.3). These price reductions reduce significantly the revenue of the aggregator from its inflexible EV, which charge during peak hours; as a result, the overall profit of the aggregator is substantially lower under DQN (Table 5.2).

5.5.5 Computational Performance

Beyond the economic performance of the three RL methods, it is highly valuable to compare their computational performance. In this context, Table 5.4 presents the average computational time per episode, the number of episodes required to reach convergence, and the total computational time required to reach convergence, for each method in the examined scenario. The average time per episode is the lowest in Q-Learning (since this method relies on look-up tables and does not require the computationally intensive training of a DNN), higher in DQN (since this method involves training of one DNN) and the highest in the proposed PDDPG (since this method involves training of two DNNs). However, the number of episodes and the total computational time is the lowest in PDDPG due to the employment of the PER strategy, followed by DQN and Q-Learning. These results demonstrate that beyond achieving a higher average profit and a smaller standard deviation of profit with respect to the two state-of-the-art RL methods, the proposed PDDPG method also exhibits a more favorable computational performance.

Table 5.4 Computational performance of the examined RL methods

Method	**Average computational time per episode (sec)**	**Number of episodes**	**Total computational time (sec)**
Q-learning	2.16	480	1,037
DQN	3.18	300	954
PDDPG	3.41	220	750

5.5.6 Impact of EV Flexibility

Having established the superiority of the proposed PDDPG method with respect to state-of-the-art methods in the previous section, the second set of studies applies this method to different scenarios regarding the nature and extent of the EV flexibility, aiming to investigate the impacts of the flexibility on the pricing decisions and the overall profit of the aggregator as well as the costs of EV owners. Specifically, two different scenarios regarding the nature of EV flexibility -smart charging capability only (denoted by G2V in the remainder) and combined smart charging and V2G capabilities (denoted by V2G in the remainder) and two different scenarios regarding the percentage of EV being flexible (25% and 50%) are examined, along with the benchmark scenario of all EV being inflexible.

For each of these scenarios, Fig. 5.9 and 5.10 illustrate the hourly profiles of the net demand of all EV in the aggregator's portfolio (positive values indicate net charging and negative values indicate net discharging), and the retail prices offered by the aggregator, respectively. When all EV are inflexible, given that a home-charging paradigm is assumed in these studies (Section 5.5.1), most of their charging demand is concentrated during the afternoon / evening (peak) hours 13-19, when most EV owners return at their homes (Fig. 5.9). Therefore, the aggregator offers the highest allowable

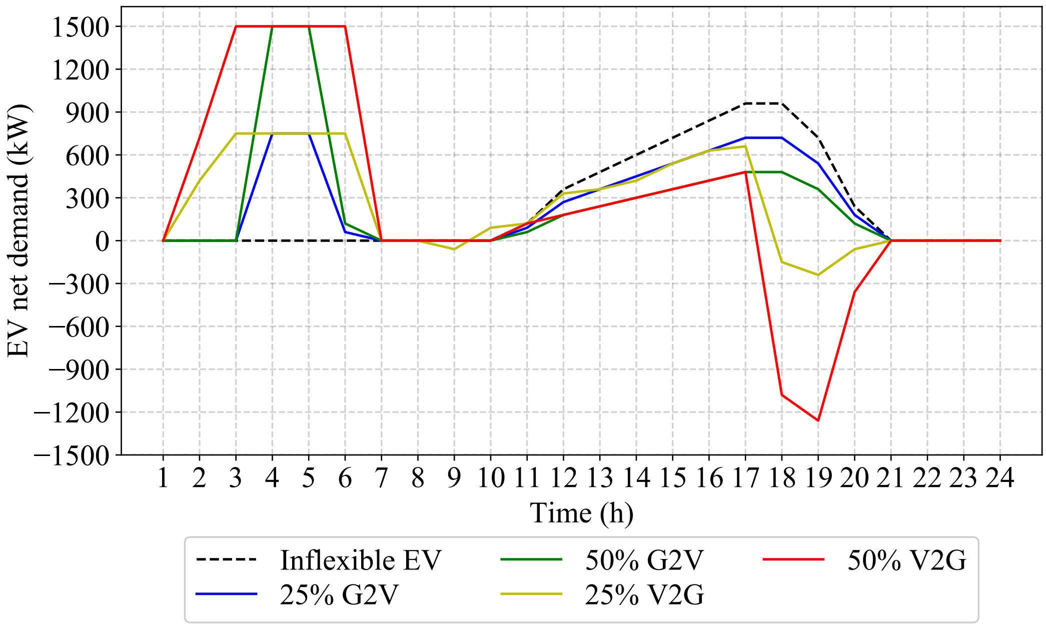

Fig. 5.9 EV net demand for different EV flexibility scenarios.

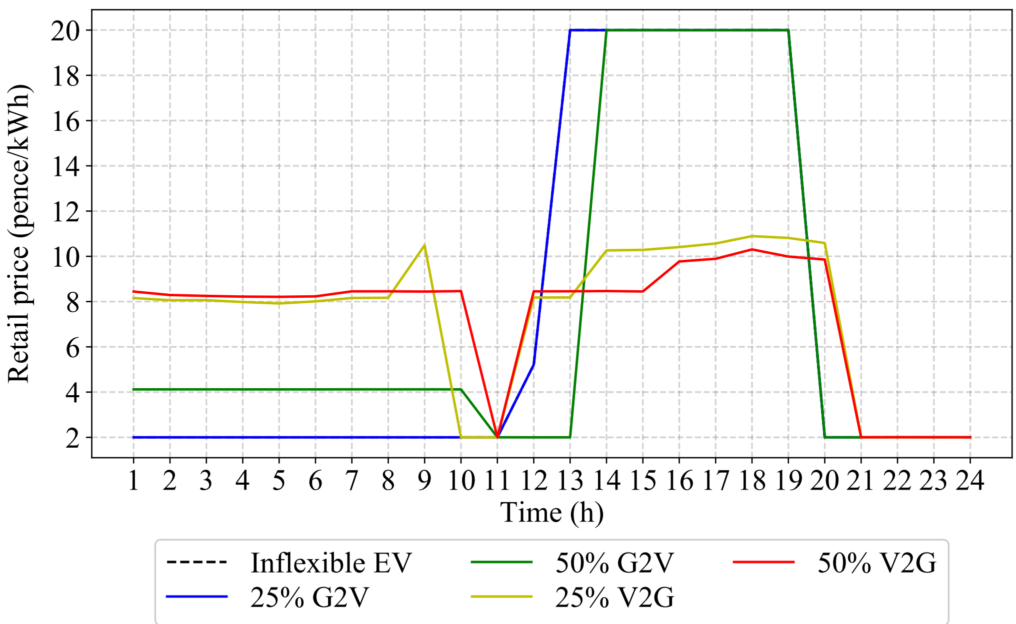

Fig. 5.10 Retail prices for different EV flexibility scenarios.

price (λ^{max} = 20 pence/kWh) at these peak hours and the lowest allowable price (λ^{min} = 2 pence/kWh) at other (off-peak) hours (Fig. 5.10) in order to maximize its revenue, which is determined by the summation of the EV demand-retail price products across all hours, but also satisfy the regulatory constraint imposed on the average retail price (5.3). In other words, the aggregator exploits the EV owners by setting large retail price differentials between peak and off-peak hours.

When EV exhibit smart charging capability, they are able to respond to the time-differentiated prices offered by the aggregator and shift their charging demand from high-price, peak hours to low-price, off-peak hours in order to minimize their cost, with this effect being enhanced as the percentage of smart charging EV increases (Fig. 5.9). While in the scenario where this percentage is lower (25% G2V) the offered prices do not substantially change (the black dashed line coincides with the blue solid line in Fig. 5.10), in the scenario where this percentage is higher (50% G2V), in anticipation of the significant shifting response, the aggregator increases its offered prices at off-peak hours in order to maintain its revenue at the highest possible level, and reduces its offered prices at hours 12-13 in order to satisfy the average retail price constraint (Fig. 5.10). In other words, a smart charging capability of significant extent results in a flatter retail price profile and thus limits the exploitation of the EV owners by the aggregator. Furthermore, in both scenarios with smart charging EV and as explained in Section 5.5.2, the aggregator offers marginally lower prices at the two hours with the lowest wholesale prices (4-5) compared to other, adjacent off-peak hours, in order to attract the demand of smart charging EV at these two hours (Fig. 5.9) and thus reduce its cost in the wholesale market (this marginal difference is not visible in Fig. 5.10 but is demonstrated in Table 5.3).

When EV exhibit V2G capability, apart from shifting their charging demand to low-price, off-peak hours, they also discharge during high-price, peak hours (as indicated by the lower net demand with respect to the smart-charging-only scenarios, which even becomes negative in some peak hours, in Fig. 5.9) to gain high revenues from selling energy. Thus, their charging window and their overall charging energy consumption during off-peak hours are larger with respect to the smart-charging-only scenarios (Fig. 5.9). In other words, the V2G EV perform energy arbitrage. In anticipation of this response, the aggregator increases further its offered prices at off-peak hours and reduces further its offered prices at peak hours, resulting in a higher prominent flattening effect on the retail price profile with respect to the smart-charging-only scenarios (Fig. 5.10).

Having analyzed the impacts of EV flexibility on the net demand of the EV and the pricing decisions of the aggregator, the final part of this analysis lies in investigating the overall economic effects of this flexibility on both the aggregator and the EV owners. Starting from the former, Table 5.5 presents the total (daily) profit of the aggregator

(5.1) along with its net retail revenue (corresponding to the sum of the first three terms of (5.1)) and net wholesale cost (corresponding to the fourth term of (5.1)) components, for each of the examined scenarios.

The net retail revenue of the aggregator is reduced as the percentage of smart charging EV increases (by 22% in the 25% G2V scenario and 43% in the 50% G2V scenario), driven by the flattening effect on the EV net demand (Fig. 5.9) and the offered retail prices (Fig. 5.10), considering that the net retail revenue of the aggregator is determined by the summation of the EV net demand-retail price products across all hours. In scenarios where EV also exhibit V2G capability, this flattening effect is enhanced and thus the net retail revenue of the aggregator is further reduced (by 53% in the 25% V2G scenario and 57% in the 50% V2G scenario). This trend implies that EV flexibility deteriorates the business case of the aggregator, since it limits the strategic potential of exploiting the EV owners by setting large retail price differentials between peak and off-peak hours.

Going further, the net wholesale cost of the aggregator is also reduced as the percentage of smart charging EV increases (by 20% in the 25% G2V scenario and 40% in the 50% G2V scenario), due to the shift of charging demand (which constitutes the demand of the aggregator in the wholesale market) from hours with high wholesale prices to hours with low wholesale prices. In scenarios where EV also exhibit V2G capability, the net demand of the aggregator in the wholesale market during high-price, peak hours is further reduced -and even becomes negative in some peak hours (Fig. 5.9) -and thus the net wholesale cost of the aggregator is further reduced (by 51% in the 25% V2G scenario and 99% in the 50% V2G scenario). This trend implies that EV flexibility improves the business case of the aggregator, since it enables the aggregator to buy more energy at low-price hours and, in the case with V2G capability, even sell energy at high-price hours.

The impact of EV flexibility on the total profit of the aggregator (which constitutes the most important index of the aggregator's business case) is intuitively driven by the combination of its impacts on the aggregator's retail revenue and wholesale cost. Since these impacts counteract each other, i.e. EV flexibility reduces its profit by reducing its retail revenue but also increases its profit by reducing its wholesale cost, the overall impact on the aggregator's profit depends on which of these two profit components is reduced at a higher rate in each of the examined scenarios. Table 5.5 demonstrates a non-uniform trend in the examined study; while the aggregator's profit is reduced in scenarios 25% G2V, 50% G2V and 25% V2G with respect to the scenario where all EV are inflexible, it is increased in scenario 50% V2G.

Table 5.5 Profit of aggregator for different EV flexibility scenarios.

Scenario	Net retail revenue (£)	Net wholesale cost (£)	Profit (£)
Inflexible EV	1081.96	723.00	358.96
25% G2V	842.57	579.68	262.89
50% G2V	620.48	436.36	184.12
25% V2G	513.31	354.02	159.29
50% V2G	462.95	8.73	454.22

Shifting our focus to the EV owners, Table 5.6 presents the average daily cost of an inflexible EV and a flexible EV, for each of the examined scenarios; the average cost is employed as a representative economic index, due to the fact that different EV are characterized by different travelling times, grid connection times and travelling energy requirements (Section 5.5.1). The introduction of flexibility, irrespectively of its nature and extent, reduces the average cost of an EV with respect to the benchmark scenario where all EV are inflexible.

This beneficial impact of flexibility is even evident for EV which are not flexible, driven by the fact that the introduction of flexibility in other EV reduces the retail prices offered by the aggregator at peak hours (Fig. 5.10), during which inflexible EV charge. As the percentage of flexible EV increases and when V2G capability is introduced, this reduction of peak retail prices is enhanced, and thus the average cost of an inflexible EV is further reduced (Table 5.6).

However, this beneficial impact is significantly higher for EV which exhibit this flexibility, since they shift their charging demand from high-price hours to low-price hours. Interestingly however, as the percentage of flexible EV increases and when V2G capability is introduced, the retail price profile becomes flatter (Fig. 5.10), and thus the average cost reduction enjoyed by a flexible EV is diminished (Table 5.6).

Table 5.6 Average cost of EV for different EV flexibility scenarios.

Scenario	Inflexible EV average cost (£)	Flexible EV average cost (£)
Inflexible EV	1.0820	N/A
25% G2V	1.0818	0.1248
50% G2V	0.9839	0.2570
25% V2G	0.6035	0.2428
50% V2G	0.5566	0.3693

5.6 Conclusions

In the context of realizing the significant EV flexibility potential in deregulated electricity systems, this chapter has addressed the problem of effectively pricing EV by aggregators through a bi-level optimization formulation. In contrast with the existing literature, this formulation also considers the V2G capability of EV and the discrete nature of their charging / discharging levels. Considering the fundamental limitations of the traditional MPEC approach in capturing these discrete EV operating characteristics and thus effectively solving this problem, this chapter has focused on RL approaches. In this setting, motivated by the solution optimality and computational limitations of state-of-the-art RL methods, associated with their need to discretize state and / or action spaces, this chapter has proposed a novel deep reinforcement learning method, named *prioritized deep deterministic policy gradient* method (PDDPG), which sets up the problem in multi-dimensional continuous state and action spaces.

The scope of the presented case studies has been twofold. First of all, the proposed PDDPG method has been compared against the two state-of-the-art RL methods, namely Q-learning and DQN. The results have demonstrated that the proposed method achieves 31% and 13% higher profit for the examined EV aggregator than the Q-learning and DQN methods, respectively, driven by its ability to capture continuous pricing decisions. Furthermore, it exhibits lower total computational requirements, driven by the employment of the proposed prioritized experience replay (PER) strategy.

Secondly, the presented case studies have applied the proposed method to different scenarios regarding the nature and extent of the EV flexibility, in order to investigate the economic impacts of such flexibility on both the aggregator and the EV owners. The results have demonstrated that as the percentage of flexible EV increases and when V2G capability is introduced, the profile of retail prices offered by the aggregator gets flatter and the aggregator's retail revenue from the EV is reduced, implying that EV flexibility limits the aggregator's potential of exploiting the EV owners by setting large retail price differentials between peak and off-peak hours. However, EV flexibility also reduces the aggregator's cost in the wholesale market, since it enables the aggregator to buy more energy at low-price hours and, in the case with V2G capability, even sell energy at high-price hours. As a result, the overall impact of EV flexibility on the aggregator's total profit is not uniformly beneficial or detrimental, but is rather scenario-specific. On the other hand, the introduction of EV flexibility has been shown to always reduce the average electricity cost of EV. Although this beneficial impact is significantly higher for the EV which exhibit this flexibility, it is also substantial for inflexible EV, driven by the impact of EV flexibility in reducing the retail prices offered by the aggregator at peak hours.

Chapter 6

Conclusions and Future Work

6.1 Conclusion

This thesis studies the strategic retail pricing methods to the flexible customers based on game-theoretic and learning based techniques in the smart grid. It includes 1) developing a comprehensive framework for electricity retailers taking both the customer's demand response and the wholesale market clearing into consideration; 2) investigating the impact of LEM on electricity retailer's pricing strategies, its profit and the served customers' economic surplus; 3) building a learning based method for the aggregator to design effective retail prices, accounting for the discrete charging / discharging response patterns of the served EV.

We concentrate on studying the interactions between the electricity retailer / aggregator and the served customers from the following three aspects:

1. The impact of demand flexibility on an electricity retailer's pricing strategies and its business case through a bi-level optimization problem is firstly investigated in Chapter 3.

In contrast with state-of-the-art bi-level optimization models of the interaction between the retailer and its customers, this model drops the unrealistic assumption that the retailer treats wholesale market prices as exogenous, fixed parameters, and represents endogenously the wholesale market clearing process as an additional lower-level problem. Thus, this chapter can be seen as our first attempt to solve the demand response based strategic retail pricing problem by capturing the realistic implications of the retailer's pricing strategies on the wholesale market prices. As the lower-level problems considered in this chapter are both continuous and convex, we proposed a

KKT condition based solution to solve the proposed bi-level optimization problem model effectively.

The presented case studies provide numerous new and valuable insights. First of all, they demonstrate that demand flexibility reduces the retailer's revenue from the consumers, since it limits the retailer's strategic potential of exploiting the consumers through setting large retail price differentials between peak and off-peak hours; and also the retailer's cost in the wholesale market, since it enables the retailer to buy more energy at low-price hours and also reduce the wholesale prices at high-price hours.

Going further, this impact of demand flexibility on the retailer's profit implies that demand flexibility can effectively complement regulatory policies in safeguarding the consumers against the strategic behavior of retailers. Specifically, under a looser regulatory framework demand flexibility reduces the overall profit of the retailer, while under a stricter regulatory framework it increases this profit.

Moreover, the presented case studies demonstrate that under relatively high demand flexibility the retailer achieves a higher profit when its size is smaller, while under relatively low demand flexibility the retailer achieves a higher profit when its size is larger. This result implies that new, small players in the retail market are more likely to take initiatives towards the realization of the flexibility potential of their consumers, than large, incumbent retailers.

Finally, the presented case studies highlight the added value of the proposed bi-level model by comparing its outcomes against the state-of-the-art bi-level model neglecting the wholesale market clearing. The result implies that the state-of-the-art model is suitable for driving a retailer's decision-making only under the limiting condition that the retailer's size is extremely small (around 1% of the market according to the obtained results), in contrast with the general suitability of the proposed model.

2. The impact of LEM on an electricity retailer's pricing strategies, its business case and the different customers' economic surplus through a bi-level optimization problem is firstly investigated in Chapter 4.

This chapter has explored for the first time the interactions between the operation of LEM with different types of participants (FC, MG and ES) and the strategic pricing decisions of incumbent electricity retailers, and has quantitatively analyzed the overall economic effects of LEM on both the retailer and its customers.

In order to achieve that, this chapter has developed a novel multi-period bi-level optimization model, which captures the pricing decisions of a retailer in the upper-level problem and the response of both independent customers and the LEM in the lower-level problems. Since the lower-level problem representing the LEM is non-convex, the traditional MPEC approach is not applicable for solving the developed bi-level problem, and a new approach recently proposed by the authors is employed instead, which is based on the relaxation and primal-dual reformulation of the non-convex lower-level problem and the penalization of the associated duality gap.

The presented case studies have provided numerous new and valuable insights around the role and impact of LEM. First of all, the introduction of an LEM is shown to reduce the customers' energy dependency on the retailer, since they are able to trade energy among them at prices which lie between the retailer's high buy prices and low sell prices, which is mutually beneficial for all FC, MG and ES participants; regarding the latter, the LEM is shown to unlock their arbitrage potential and activate them in the market. As a consequence, the retailer's strategic potential of exploiting the customers through large differentials between buy and sell prices is limited, and the retailer strives to make its offered buy and sell prices more competitive in order to attract more demand and generation by its customers. As a result of these effects, the profit of the retailer is very significantly reduced, while the customers enjoy significant economic benefits. Although this beneficial impact of LEM is significantly higher for customers participating in the LEM, it is also substantial for non-participating customers, due to the above effects of the LEM on the retailer's offered prices.

3. A novel deep reinforcement learning method to solve the pricing EV problem with discrete charging / discharging levels is firstly proposed in Chapter 5.

This chapter solves the problem of effectively pricing EV by aggregators through a bi-level optimization formulation. In contrast with the existing literature, this formulation also considers the V2G capability of EV and the discrete nature of their charging / discharging levels.

Considering the fundamental limitations of the traditional KKT condition based approach in capturing these discrete EV operating characteristics and thus effectively solving this problem, this chapter has focused on RL approaches. In this setting, motivated by the solution optimality and computational limitations of state-of-the-art RL methods, associated with their need to discretize state and / or action spaces, this chapter has proposed a novel deep reinforcement learning method, named prioritized deep deterministic policy gradient method (PPDPG), which sets up the problem in multi-dimensional continuous state and action spaces.

The presented case studies firstly demonstrate that the proposed PDDPG method has been compared against the two state-of-the-art RL methods, namely Q-learning and

DQN. The results have demonstrated that the proposed method achieves 31% and 13% higher profit for the examined EV aggregator than the Q-learning and DQN methods, respectively, driven by its ability to capture continuous pricing decisions. Furthermore, it exhibits lower total computational requirements, driven by the employment of the proposed prioritized experience replay (PER) strategy.

The presented case studies secondly demonstrated that as the percentage of flexible EV increases and when V2G capability is introduced, the profile of retail prices offered by the aggregator gets flatter and the aggregator's retail revenue from the EV is reduced, implying that EV flexibility limits the aggregator's potential of exploiting the EV owners by setting large retail price differentials between peak and off-peak hours.

On the other hand, the introduction of EV flexibility has been shown to always reduce the average electricity cost of EV. Although this beneficial impact is significantly higher for the EV which exhibit this flexibility, it is also substantial for inflexible EV, driven by the impact of EV flexibility in reducing the retail prices offered by the aggregator at peak hours.

6.2 Future Work

Although the thesis fulfils the aims of developing efficient smart pricing strategies to flexible customers in the retail market via game-theoretic and learning based techniques, there is still some work that can be developed in the future.

1. The developed model in Chapter 3 as well as the similar bi-level optimization models employed in the existing literature for analysing strategic decision-making in deregulated electricity markets, neglect the complex unit commitment constraints of electricity producers in the wholesale market. This limitation is associated with the mathematical challenge of rigorously solving a bi-level optimization problem with binary decision variables in the lower-level problem, given that the derivation of the respective KKT optimality conditions is impossible. However, these complex constraints generally affect the wholesale market outcome and subsequently the retailer's pricing strategies and profit. In this context, future work will enable the incorporation of binary unit commitment variables in the developed model through 1) the mathematical approach of the relaxation and primal-dual reformulation of the non-convex lower-level problem and the penalization of the associated duality gap, which is adopted in Chapter 4 or 2) the reinforcement learning approach of learning the strategies by acquiring the

experiences from the environment that can be modeled as a non-convex optimization problem, which is adopted in Chapter 5.

2. Beyond the generic, technology-agnostic representation of demand flexibility employed in Chapter 3 and 4, the detailed representation of different residential and commercial flexible demand technologies, including electric vehicles [34], electric heat pumps [30] and smart appliances [35], will be incorporated in the model.

3. The network is also an important issue in the power system. Our future work will consider how the network constraints can be integrated into the developed optimization framework in Chapter 3-5 and see the effects of the line congestion on the market outcomes and the strategic decisions of the retailer as well as its businesses. Take the bi-level model of Chapter 3 as an example, in order to consider the effect of the transmission network, the DC power flow model proposed in [89, 131] will be implemented to the second lower-level problem (i.e. wholesale market clearing); while for the effect of the distribution network, the AC power flow model proposed in [26] could be implemented to the upper-level problem. And the approximated network equations developed in [132] are necessary to keep the model linear.

4. The additional one lies in capturing the strategic interaction between multiple independent retailers by extending the proposed model to an equilibrium programming model determining retail market equilibria and considering the ability of consumers to switch electricity retailer, depending on the offered retail prices. Authors in [133–136] formulate this equilibrium problem by replacing each player's MPEC problem by its KKT optimality conditions and concatenate them together, resulting in a set of nonlinear constraints known as *equilibrium problem with equilibrium constraints* (EPEC). An *iterative diagonalization algorithm* (DIAG) is used in [89, 131, 137, 138] to identify the imperfect equilibrium, in which each retailer solves its own MPEC problem treating the strategies of the rest of the retailers as fixed, until the algorithm converges to a fixed market outcome. Furthermore, *multi-agent reinforcement learning* (MARL) is also adopted to address the equilibrium problem in the context of reinforcement learning algorithm [139, 140].

5. Furthermore, the presented models in Chapter 3-5 are deterministic, assuming that the examined retailer has accurate projections of its customers' flexibility and the generation / demand characteristics of the wholesale market. Future work aims at incorporating uncertainties that retailers face regarding these parameters and investigating the retailers' strategies and business case in this setting, rendering the problem into a stochastic optimization problem. In addition, the retailer also accounts for potential hedging strategies through their participation in financial derivative

markets [141], such as forward contacts, futures, swaps, and options, that are introduced to manage the risk coming from the uncertainties of wholesale prices and consumers' flexibility, because the prices of these derivatives are predefined, but normally are expensive than the wholesale prices. More specifically, a risk-taken retailer is willing to accept the risk from wholesale prices, and thus purchases more energy from the wholesale market and less energy from financial derivatives, so as to reduce its expected cost and increase its expected profit. While a risk-aversion retailer aims at mitigating the risk from wholesale prices, and thus reduces the energy procurement from the wholesale market and increases the energy procurement from financial derivatives [142]. However, its expected profit is reduced because the risk-aversion retailer purchases more expensive financial derivatives.

6. Future work aims at enhancing the developed model in Chapter 4-5. The developed model assumes that the retailer / aggregator's decisions do not affect the wholesale market prices, which constitutes a realistic assumption only for retailer / aggregator serving a small population of customers / EV. Future work aims at dropping this assumption and exploring the effects of the introduction of LEM / EV flexibility on large retailer / aggregator who can also act strategically in the wholesale market, but also considers the more realistic constraints incorporating the model dynamics.

7. Specifically, for Chapter 4, the future work aims at model the optimal decisions of the customers to choose between the strategic retailer and LEM. To this end, it would be interesting to implement such a complex model in a multi-agent deep reinforcement learning (MADRL) framework. The strategic retailer and the served customers are both treated as the agents, where the retailer aims at maximizing its profit (reward), receives the response from individual customers and LEM (state) and offers the strategic prices (action); the customers aim at maximizing their individual benefits (reward), receives the price signals offered by the retailer (state) and optimize their selection choice and optimal response (action).

8. Finally, the future work of Chapter 5 aims at incorporating the realistic variability of the exogenous state features (namely EV traveling patterns and wholesale prices) in the learning procedure of the proposed PDDPG method, through the employment of data from large EV trials. This will help us test and enhance the generalization performance of the proposed method to render it robust against such variability, which constitutes a major current research challenge in the area of RL.

References

[1] European Commission, 2030 climate & energy framework. [Online]. Available: https://ec.europa.eu/clima/policies/strategies/2030_en

[2] G. Strbac, "Demand side management: Benefits and challenges," *Energy policy*, vol. 36, no. 12, pp. 4419–4426, Dec. 2008.

[3] C.-L. Su and D. Kirschen, "Quantifying the effect of demand response on electricity markets," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1199–1207, Jul. 2009.

[4] C. De Jonghe, B. F. Hobbs, and R. Belmans, "Optimal generation mix with short-term demand response and wind penetration," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 830–839, Jan. 2012.

[5] D. Pudjianto and G. Strbac, "Assessing the value and impact of demand side response using whole-system approach," *Proc. IMechE Part A: J Power and Energy*, vol. 231, no. 6, pp. 498–507, Sep. 2017.

[6] D. Papadaskalopoulos, R. Moreira, G. Strbac, D. Pudjianto, P. Djapic, F. Teng, and M. Papapetrou, "Quantifying the potential economic benefits of flexible industrial demand in the european power system," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 5123–5132, Mar. 2018.

[7] D. S. Kirschen and G. Strbac, *Fundamentals of power system economics.* John Wiley & Sons, 2018.

[8] J. Yang, J. Zhao, F. Luo, F. Wen, and Z. Y. Dong, "Decision-making for electricity retailers: A brief survey," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4140–4153, Jan. 2017.

[9] Electricity supply market shares by company: Domestic (GB), 2017. [Online]. Available: https://www.ofgem.gov.uk/data-portal/electricity-supply-market-shares-company-domestic-gb

[10] L. Hongling, J. Chuanwen, and Z. Yan, "A review on risk-constrained hydropower scheduling in deregulated power market," *Renew. Sustain. Energy Rev.*, vol. 12, no. 5, pp. 1465–1475, Jun. 2008.

[11] Y. Liu and F. F. Wu, "Generator bidding in oligopolistic electricity markets using optimal control: fundamentals and application," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1050–1061, Aug. 2006.

[12] T. Krause, E. V. Beck, R. Cherkaoui, A. Germond, G. Andersson, and D. Ernst, "A comparison of nash equilibria analysis and agent-based modelling for power markets," *Int. J. Elec. Power*, vol. 28, no. 9, pp. 599–607, Nov. 2006.

[13] D. S. Kirschen, "Demand-side view of electricity markets," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 520–527, May 2003.

[14] R. Deng, Z. Yang, M.-Y. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 570–582, Jun. 2015.

[15] W. Saad, A. L. Glass, N. B. Mandayam, and H. V. Poor, "Toward a consumer-centric grid: A behavioral perspective," *Proc. IEEE*, vol. 104, no. 4, pp. 865–882, May 2016.

[16] W. Su, "The role of customers in the us electricity market: Past, present and future," *Electr. J.*, vol. 27, no. 7, pp. 112–125, Aug. 2014.

[17] T. Chen, H. Pourbabak, Z. Liang, and W. Su, "An integrated evoucher mechanism for flexible loads in real-time retail electricity market," *IEEE Access*, vol. 5, pp. 2101–2110, Jan. 2017.

[18] T. Chen, H. Pourbabak, and W. Su, "A game theoretic approach to analyze the dynamic interactions of multiple residential prosumers considering power flow constraints," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1–5.

[19] Z. Xu, D. S. Callaway, Z. Hu, and Y. Song, "Hierarchical coordination of heterogeneous flexible loads," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4206–4216, Jan. 2016.

[20] D. S. Callaway and I. A. Hiskens, "Achieving controllability of electric loads," *Proc. IEEE*, vol. 99, no. 1, pp. 184–199, Nov. 2010.

[21] W. Su and M.-Y. Chow, "Performance evaluation of an eda-based large-scale plug-in hybrid electric vehicle charging algorithm," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 308–315, Jun. 2011.

[22] J. Yusta, I. Ramirez-Rosado, J. Dominguez-Navarro, and J. Perez-Vidal, "Optimal electricity price calculation model for retailers in a deregulated market," *Int. J. Elect. Power Energy Syst.*, vol. 27, no. 5-6, pp. 437–447, Jun. 2005.

[23] P. Faria and Z. Vale, "Demand response in electrical energy supply: An optimal real time pricing approach," *Energy*, vol. 36, no. 8, pp. 5374–5384, Aug. 2011.

[24] M. Doostizadeh and H. Ghasemi, "A day-ahead electricity pricing model based on smart metering and demand-side management," *Energy*, vol. 46, no. 1, pp. 221–230, Otc. 2012.

[25] R. García-Bertrand, "Sale prices setting tool for retailers," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2028–2035, May 2013.

[26] A. Safdarian, M. Fotuhi-Firuzabad, and M. Lehtonen, “Integration of price-based demand response in discos’ short-term decision model,” *IEEE Trans. Smart Grid*, vol. 5, no. 5, pp. 2235–2245, Jul. 2014.

[27] S. Nojavan, K. Zare, and B. Mohammadi-Ivatloo, “Optimal stochastic energy management of retailer based on selling price determination under smart grid environment in the presence of demand response program,” *Appl. Energy*, vol. 187, pp. 449–464, Feb. 2017.

[28] S. Nojavan and K. Zare, “Optimal energy pricing for consumers by electricity retailer,” *Electr. Power Energy Syst.*, vol. 102, pp. 401–412, Nov. 2018.

[29] F. Hu, X. Feng, and H. Cao, “A short-term decision model for electricity retailers: Electricity procurement and time-of-use pricing,” *Energies*, vol. 11, no. 12, p. 3258, Nov. 2018.

[30] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, “A bilevel model for electricity retailers’ participation in a demand response market environment,” *Energy Econ.*, vol. 36, pp. 182–197, May 2013.

[31] W. Wei, F. Liu, and S. Mei, “Energy pricing and dispatch for smart grid retailers under demand response and market price uncertainty,” *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1364–1374, Dec. 2014.

[32] S. Bahramara, M. P. Moghaddam, and M. R. Haghifam, “Modelling hierarchical decision making framework for operation of active distribution grids,” *IET Gener. Trans. Distrib.*, vol. 9, no. 16, pp. 2555–2564, Dec. 2015.

[33] S. Bahramara, M. P. Moghaddam, and M. Haghifam, “A bi-level optimization model for operation of distribution networks with micro-grids,” *Electr. Power Energy Syst.*, vol. 82, pp. 169–178, Nov. 2016.

[34] I. Momber, S. Wogrin, and T. G. San Román, “Retail pricing: A bilevel program for pev aggregator decisions using indirect load control,” *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 464–473, Feb. 2015.

[35] D. T. Nguyen, H. T. Nguyen, and L. B. Le, “Dynamic pricing design for demand response integration in power distribution networks,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3457–3472, Jan. 2016.

[36] M. Yu and S. H. Hong, “Supply–demand balancing for power management in smart grid: A stackelberg game approach,” *Appl. energy*, vol. 164, pp. 702–710, Feb. 2016.

[37] S. Sekizaki, I. Nishizaki, and T. Hayashida, “Electricity retail market model with flexible price settings and elastic price-based demand responses by consumers in distribution network,” *Int. J. Electr. Power Energy Syst.*, vol. 81, pp. 371–386, Otc. 2016.

[38] J. Yang, J. Zhao, F. Wen, and Z. Y. Dong, “A framework of customizing electricity retail prices,” *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2415–2428, Sep. 2017.

[39] R. Lu, S. H. Hong, and X. Zhang, "A dynamic pricing demand response algorithm for smart grid: reinforcement learning approach," *Appl. Energy*, vol. 220, pp. 220–230, Jun. 2018.

[40] H. Karimi, S. Jadid, and H. Saboori, "Multi-objective bi-level optimisation to design real-time pricing for demand response programs in retail markets," *IET Gener. Transm. Distrib.*, vol. 13, no. 8, pp. 1287–1296, Dec. 2018.

[41] A.-Y. Yoon, Y.-J. Kim, and S.-I. Moon, "Optimal retail pricing for demand response of hvac systems in commercial buildings considering distribution network voltages," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5492–5505, Nov. 2018.

[42] A. J. Conejo, M. Carrión, J. M. Morales *et al.*, *Decision making under uncertainty in electricity markets.* Springer, 2010, vol. 1.

[43] E. Celebi and J. D. Fuller, "Time-of-use pricing in electricity markets under different market structures," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1170–1181, Feb. 2012.

[44] P. Warren, "A review of demand-side management policy in the uk," *Renew. Sustain. Energy Rev.*, vol. 29, pp. 941–951, Jan. 2014.

[45] M. Behrangrad, "A review of demand side management business models in the electricity market," *Renew. Sustain. Energy Rev.*, vol. 47, pp. 270–283, Jul. 2015.

[46] M. F. Akorede, H. Hizam, and E. Pouresmaeil, "Distributed energy resources and benefits to the environment," *Renew. Sustain. Energy Rev.*, vol. 14, no. 2, pp. 724–734, Feb. 2010.

[47] D. Wu, T. Yang, A. A. Stoorvogel, and J. Stoustrup, "Distributed optimal coordination for distributed energy resources in power systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 414–424, Dec. 2016.

[48] M. Carrión and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1371–1378, Jul. 2006.

[49] J. S. Netz, "Price regulation: A (non-technical) overview," *Encycl. Law Econ.*, vol. 3, pp. 396–466, 2000.

[50] M. P. Moghaddam, A. Abdollahi, and M. Rashidinejad, "Flexible demand response programs modeling in competitive electricity markets," *Appl. Energy*, vol. 88, no. 9, pp. 3257–3269, Sep. 2011.

[51] M. Aunedi, "Value of flexible demand-side technologies in future low-carbon systems," Ph.D. dissertation, Imperial College London, 2013.

[52] D. Papadaskalopoulos, "A mechanism for decentralized participation of flexible demand in electricity markets," Ph.D. dissertation, Imperial College London, 2013.

[53] E. Bompard, Y. Ma, R. Napoli, and G. Abrate, "The demand elasticity impacts on the strategic bidding behavior of the electricity producers," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 188–197, Jan. 2007.

[54] E. Bompard, Y. Ma, R. Napoli, G. Abrate, and E. Ragazzi, "The impacts of price responsiveness on strategic equilibrium in competitive electricity markets," *Elec. Power Ener. Syst.*, vol. 29, no. 5, pp. 397–407, Jun. 2007.

[55] E. Bompard, R. Napoli, and B. Wan, "The effect of the programs for demand response incentives in competitive electricity markets," *Eur. Trans. Elec. Power*, vol. 19, no. 1, pp. 127–139, Jan. 2009.

[56] P. R. Thimmapuram, J. Kim, A. Botterud, and Y. Nam, "Modeling and simulation of price elasticity of demand using an agent-based model," in *Innov. Smart Grid Tech. (ISGT)*, 2010, pp. 1–8.

[57] D. Pozo and J. Contreras, "Finding multiple nash equilibria in pool-based markets: A stochastic epec approach," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1744–1752, Jan. 2011.

[58] S. Chou, W. Yang, K. Chua, J. Li, and K. Zhang, "Development of micro power generators–a review," *Appl. Energy*, vol. 88, no. 1, pp. 1–16, Jan. 2011.

[59] S. Allen, G. Hammond, and M. McManus, "Prospects for and barriers to domestic micro-generation: A united kingdom perspective," *Appl. Energy*, vol. 85, no. 6, pp. 528–544, Jun. 2008.

[60] D. Infield and F. Li, "Integrating micro-generation into distribution systems—a review of recent research," in *Proc. 2008 IEEE Power and Energy Soc. General Meeting—Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1–4.

[61] J. Castaneda, J. Enslin, D. Elizondo, N. Abed, and S. Teleke, "Application of statcom with energy storage for wind farm integration," in *IEEE PES T&D 2010*, 2010, pp. 1–6.

[62] B. P. Roberts and C. Sandberg, "The role of energy storage in development of smart grids," *Proc. IEEE*, vol. 99, no. 6, pp. 1139–1144, Jun. 2011.

[63] C. Liu, K. Chau, D. Wu, and S. Gao, "Opportunities and challenges of vehicle-to-home, vehicle-to-vehicle, and vehicle-to-grid technologies," *Proc. IEEE*, vol. 101, no. 11, pp. 2409–2427, Jul. 2013.

[64] W. Tushar, W. Saad, H. V. Poor, and D. B. Smith, "Economics of electric vehicle charging: A game theoretic approach," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1767–1778, Sep. 2012.

[65] B.-G. Kim, S. Ren, M. Van Der Schaar, and J.-W. Lee, "Bidirectional energy trading and residential load scheduling with electric vehicles in the smart grid," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1219–1234, Jun. 2013.

[66] A. Y. Lam, L. Huang, A. Silva, and W. Saad, “A multi-layer market for vehicle-to-grid energy trading in the smart grid,” in *2012 Proceedings IEEE INFOCOM Workshops*, 2012, pp. 85–90.

[67] C. Wu, H. Mohsenian-Rad, and J. Huang, “Vehicle-to-aggregator interaction game,” *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 434–442, Oct. 2011.

[68] E. Sortomme and M. A. El-Sharkawi, “Optimal combined bidding of vehicle-to-grid ancillary services,” *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 70–79, Nov. 2011.

[69] J. Zheng, X. Wang, K. Men, C. Zhu, and S. Zhu, “Aggregation model-based optimization for electric vehicle charging strategy,” *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 1058–1066, Jun. 2013.

[70] I. S. Bayram, M. Z. Shakir, M. Abdallah, and K. Qaraqe, “A survey on energy trading in smart grid,” in *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2014, pp. 258–262.

[71] D. Niyato, P. Wang, and E. Hossain, “Reliability analysis and redundancy design of smart grid wireless communications system for demand side management,” *IEEE Wireless Communications*, vol. 19, no. 3, pp. 38–46, Jul. 2012.

[72] J. Wang, “Conjectural variation-based bidding strategies with q-learning in electricity markets,” in *2009 42nd Hawaii International Conference on System Sciences*, 2009, pp. 1–10.

[73] D. Pozo, E. Sauma, and J. Contreras, “Basic theoretical foundations and insights on bilevel models and their applications to power systems,” *Ann. Oper. Res.*, vol. 254, no. 1-2, pp. 303–334, Jul. 2017.

[74] J. F. Bard, *Practical bilevel optimization: algorithms and applications.* Springer Science & Business Media, 2013, vol. 30.

[75] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, *Complementarity modeling in energy markets.* Springer Science & Business Media, 2012, vol. 180.

[76] B. Colson, P. Marcotte, and G. Savard, “An overview of bilevel optimization,” *Ann. Oper. Res.*, vol. 153, no. 1, pp. 235–256, Sep. 2007.

[77] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[78] J. T. Moore and J. F. Bard, “The mixed integer linear bilevel programming problem,” *Oper. Res.*, vol. 38, no. 5, pp. 911–921, Oct. 1990.

[79] S. Pineda and J. M. Morales, “Solving linear bilevel problems using big-ms: not all that glitters is gold,” *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2469–2471, Jan. 2019.

[80] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction.* MIT press, 2018.

[81] M. Minsky, "Steps toward artificial intelligence," *Proc. IRE*, vol. 49, no. 1, pp. 8–30, Jan. 1961.

[82] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized hamilton-jacobi-bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, Dec. 1997.

[83] C. J. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, no. 3-4, pp. 279–292, May 1992.

[84] G. Tesauro, "Temporal difference learning and td-gammon," *Communications of the ACM*, vol. 38, no. 3, pp. 58–68, Mar. 1995.

[85] S. Thrun, "Learning to play the game of chess," in *Advances in neural information processing systems*, 1995, pp. 1069–1076.

[86] V. Mnih and *et al.*, "Playing atari with deep reinforcement learning," in *Proc. NIPS'13 Deep Learn. Workshop*, Lake Tahoe, USA, Dec. 2013, pp. 1–9.

[87] ——, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[88] T. P. Lillicrap and *et al.*, "Continuous control with deep reinforcement learning," in *Proc. 4th Int. Conf. Learn. Represent. (ICLR)*, San Juan, USA, 2016, pp. 1–14.

[89] Y. Ye, D. Papadaskalopoulos, and G. Strbac, "Investigating the ability of demand shifting to mitigate electricity producers' market power," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3800–3811, Dec. 2017.

[90] A. G. Bakirtzis, N. P. Ziogos, A. C. Tellidou, and G. A. Bakirtzis, "Electricity producer offering strategies in day-ahead energy market with step-wise offers," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1804–1818, Otc. 2007.

[91] C. Ruiz and A. J. Conejo, "Pool strategy of a producer with endogenous formation of locational marginal prices," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1855–1866, Sep. 2009.

[92] J. Fortuny-Amat and B. McCarl, "A representation and economic interpretation of a two-level programming problem," *J. Oper. Res. Soc.*, vol. 32, no. 9, pp. 783–792, Sep. 1981.

[93] FICO XPRESS website. [Online]. Available: http://www.fico.com/en/Products/DMTools/Pages/FICO-Xpress-Optimization-Suite.aspx.

[94] A. Shakoor, G. Davies, and G. Strbac, "Roadmap for flexibility services to 2030," *A report to the Committee on Climate Change. London: Pöyry*, May 2017.

[95] I. J. Perez-Arriaga, “The transmission of the future: The impact of distributed energy resources on the network,” *IEEE Power and Energy Mag.*, vol. 14, no. 4, pp. 41–53, Jun. 2016.

[96] D. Pudjianto, M. Aunedi, P. Djapic, and G. Strbac, “Whole-systems assessment of the value of energy storage in low-carbon electricity systems,” *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 1098–1109, Sep. 2013.

[97] D. Papadaskalopoulos, D. Pudjianto, and G. Strbac, “Decentralized coordination of microgrids with flexible demand and energy storage,” *IEEE Trans. Sustain. Energy*, vol. 5, no. 4, pp. 1406–1414, Apr. 2014.

[98] European Commission, “Clean Energy for All Europeans”. [Online]. Available: https://ec.europa.eu/energy/en/topics/energy-strategy-and-energy-union/clean-energy-all-europeans

[99] M. Gancheva, S. O’Brien, N. Crook, and C. Monteiro, “Models of local energy ownership and the role of local energy communities in energy transition in europe,” *European Committee of the Regions. Commission for the Environment*, 2018.

[100] F. Teotia and R. Bhakar, “Local energy markets: Concept, design and operation,” in *Proc. 2016 Nat. Power Syst. Conf. (NPSC)*, 2016, pp. 1–6.

[101] E. Mengelkamp, B. Notheisen, C. Beer, D. Dauer, and C. Weinhardt, “A blockchain-based smart grid: towards sustainable local energy markets,” *Comput. Sci. Res. Dev.*, vol. 33, no. 1-2, pp. 207–214, Feb. 2018.

[102] J. M. López-Lezama, A. Padilha-Feltrin, J. Contreras, and J. I. Muñoz, “Optimal contract pricing of distributed generation in distribution networks,” *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 128–136, May 2010.

[103] G. E. Asimakopoulou and N. D. Hatziargyriou, “Evaluation of economic benefits of der aggregation,” *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 499–510, Aug. 2017.

[104] Y. Ye, D. Papadaskalopoulos, J. Kazempour, and G. Strbac, “Incorporating non-convex operating characteristics into bi-level optimization electricity market models,” *IEEE Trans. Power Syst.*, Jun. 2019.

[105] D. J. White and G. Anandalingam, “A penalty function approach for solving bi-level linear programs,” *Journal of Global Optimization*, vol. 3, no. 4, pp. 397–419, Dec. 1993.

[106] F. E. Torres, “Linearization of mixed-integer products,” *Math. Program.*, vol. 49, no. 1, pp. 427–428, Nov. 1990.

[107] M. V. Pereira, S. Granville, M. H. Fampa, R. Dix, and L. A. Barroso, “Strategic bidding under uncertainty: a binary expansion approach,” *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 180–188, Feb. 2005.

[108] L. A. Barroso, R. D. Carneiro, S. Granville, M. V. Pereira, and M. H. Fampa, "Nash equilibrium in strategic bidding: A binary expansion approach," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 629–638, May 2006.

[109] W. Kempton and J. Tomić, "Vehicle-to-grid power fundamentals: Calculating capacity and net revenue," *J. Power Sources*, vol. 144, no. 1, pp. 268–279, Jun. 2005.

[110] J. A. P. Lopes, F. J. Soares, and P. M. R. Almeida, "Integration of electric vehicles in the electric power system," *Proc. of the IEEE*, vol. 99, no. 1, pp. 168–183, Jan. 2011.

[111] P. J. Ramírez, D. Papadaskalopoulos, and G. Strbac, "Co-optimization of generation expansion planning and electric vehicles flexibility," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1609–1619, May 2015.

[112] S. Burger, J. P. Chaves-Ávila, C. Batlle, and I. J. Pérez-Arriaga, "A review of the value of aggregators in electricity systems," *Renew. Sustainable Energy Rev.*, vol. 77, pp. 395–405, Apr. 2017.

[113] G. Strbac and *et al.*, "Cost-effective decarbonization in a decentralized market: The benefits of using flexible technologies and resources," *IEEE Power Ener. Mag.*, vol. 17, no. 2, pp. 25–36, Mar.-Apr. 2019.

[114] H. Rashidizadeh-Kermani, M. Vahedipour-Dahraie, H. Najafi, A. Anvari-Moghaddam, and J. Guerrero, "A stochastic bi-level scheduling approach for the participation of ev aggregators in competitive electricity markets," *Appl. Sci.*, vol. 7, no. 10, pp. 1–16, Otc. 2017.

[115] H. Rashidizadeh-Kermani, H. Najafi, A. Anvari-Moghaddam, and J. Guerrero, "Optimal decision-making strategy of an electric vehicle aggregator in short-term electricity markets," *Energies*, vol. 11, no. 9, pp. 1–20, Sep. 2018.

[116] D. B. Richardson, "Electric vehicles and the electric grid: A review of modeling approaches, impacts, and renewable energy integration," *Renew. Sustain. Energy Rev*, vol. 19, pp. 247–254, Mar. 2013.

[117] L. Gan, U. Topcu, and S. H. Low, "Stochastic distributed protocol for electric vehicle charging with discrete charging rate," in *Proc. IEEE Power Energy Soc. Gen. Meet.*, San Diego, CA, USA, 2012, pp. 1–8.

[118] G. Binetti, A. Davoudi, D. Naso, B. Turchiano, and F. L. Lewis, "Scalable real-time electric vehicles charging with discrete charging rates," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2211–2220, Sep. 2015.

[119] B. Sun, Z. Huang, X. Tan, and D. H. Tsang, "Optimal scheduling for electric vehicle charging with discrete charging levels in distribution grid," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 624–634, Mar. 2018.

[120] M. Peters, W. Ketter, M. Saar-Tsechansky, and J. Collins, "A reinforcement learning approach to autonomous decision-making in smart electricity markets," *Mach. Learn.*, vol. 92, no. 1, pp. 5–39, Apr. 2013.

[121] B.-G. Kim, Y. Zhang, M. Van Der Schaar, and J.-W. Lee, "Dynamic pricing and energy consumption scheduling with reinforcement learning," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2187–2198, Sep. 2016.

[122] B. J. Claessens, P. Vrancx, and F. Ruelens, "Convolutional neural networks for automatic state-time feature extraction in reinforcement learning applied to residential load control," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3259–3269, Jul. 2018.

[123] E. Mocanu and *et al.*, "On-line building energy optimization using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3698–3708, Jul. 2019.

[124] Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-free real-time ev charging scheduling based on deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5246–5257, Sep. 2019.

[125] H. Wang and Y. Yu, "Exploring multi-action relationship in reinforcement learning," in *14th PRICAI Conf.*, 2016, pp. 574–587.

[126] R. Bellman, "On the theory of dynamic programming," *Proc. N. A. S.*, vol. 38, no. 8, pp. 716–719, Aug. 1952.

[127] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," in *Proc. 4th Int. Conf. Learn. Represent. (ICLR)*, San Juan, US, 2016, pp. 1–21.

[128] National Travel Survey, Department for Transport, U.K., 2008. [Online]. Available: http://www.dft.gov.uk

[129] M. Abadi and *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[130] "Xpress optimizer python interface," 2017. [Online]. Available: https://www.msi-jp.com/xpress/learning/square/01-python-interface.pdf

[131] Y. Ye, D. Papadaskalopoulos, R. Moreira, and G. Strbac, "Investigating the impacts of price-taking and price-making energy storage in electricity markets through an equilibrium programming model," *IET Gener. Transm. Dis.*, vol. 13, no. 2, pp. 305–315, Nov. 2018.

[132] N. Alguacil, A. L. Motto, and A. J. Conejo, "Transmission expansion planning: A mixed-integer lp approach," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1070–1077, Jul. 2003.

[133] W. Xian, L. Yuzeng, and Z. Shaohua, "Oligopolistic equilibrium analysis for electricity markets: a nonlinear complementarity approach," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1348–1355, Aug. 2004.

[134] J. Yao, I. Adler, and S. S. Oren, "Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network," *Oper. Res.*, vol. 56, no. 1, pp. 34–47, Feb. 2008.

[135] C. Ruiz, A. J. Conejo, and Y. Smeers, “Equilibria in an oligopolistic electricity pool with stepwise offer curves,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 752–761, May. 2012.

[136] A. Shahmohammadi, R. Sioshansi, A. J. Conejo, and S. Afsharnia, “Market equilibria and interactions between strategic generation, wind, and storage,” *Appl. Energy*, vol. 220, pp. 876–892, Jun. 2018.

[137] C. A. Berry, B. F. Hobbs, W. A. Meroney, R. P. O’Neill, and W. R. Stewart Jr, “Analyzing strategic bidding behavior in transmission networks,” *Util. Pol.*, vol. 8, no. 3, pp. 139–158, Jan. 1999.

[138] J. D. Weber and T. J. Overbye, “An individual welfare maximization algorithm for electricity markets,” *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 590–596, Aug. 2002.

[139] M. L. Littman, “Markov games as a framework for multi-agent reinforcement learning,” in *Machine learning proceedings 1994*. Elsevier, 1994, pp. 157–163.

[140] L. Buşoniu, R. Babuška, and B. De Schutter, “Multi-agent reinforcement learning: An overview,” in *Innovations in multi-agent systems and applications-1*. Springer, 2010, pp. 183–221.

[141] S.-J. Deng and S. S. Oren, “Electricity derivatives and risk management,” *Energy*, vol. 31, no. 6-7, pp. 940–953, May 2006.

[142] A. Hatami, H. Seifi, and M. K. Sheikh-El-Eslami, “A stochastic-based decision-making framework for an electricity retailer: Time-of-use pricing and electricity portfolio optimization,” *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1808–1816, Feb. 2011.

Appendix A

Travelling Data of 25 EV Types

It is assumed that each EV makes two journeys per day, an assumption reflecting the intuitive commuting patterns, mainly involving journeys from the home to the work place and vice versa, while the EVs are assumed connected to the grid between the end of their last trip and the start of their first trip of the day, in line with the home-charging paradigm. Each type is defined by the combination of the start time, end time and travel energy requirement of each of its two daily journeys. The total number of EV types considered in the case studies of Chapter 5 is 25 and their respective characteristics are presented in Table A.1.

Each inflexible EV is assumed to start charging (with the same, fixed rate of 3 kW) immediately after it is connected to the grid until it covers its daily travelling energy requirements and their respective fix charging characteristics are presented in Table A.2.

Table A.1 Travelling characteristics of 25 EV types.

Type	1st journey			2nd journey		
	Start time (h)	End time (h)	Energy requirement (kWh)	Start time (h)	End time (h)	Energy requirement (kWh)
1	8	8	2.2350	10	10	2.1547
2	8	8	2.2350	11	11	2.1547
3	8	8	2.2350	12	12	2.1547
4	8	8	2.2350	13	13	2.1547
5	8	8	2.2350	14	14	2.1547
6	8	8	2.2350	15	15	2.1547
7	8	8	2.2350	16	16	2.1547
8	8	8	2.2350	17	17	2.1547
9	8	8	2.2350	18	18	2.1547
10	9	9	2.2350	11	11	2.1547
11	9	9	2.2350	12	12	2.1547
12	9	9	2.2350	13	13	2.1547
13	9	9	2.2350	14	14	2.1547
14	9	9	2.2350	15	15	2.1547
15	9	9	2.2350	16	16	2.1547
16	9	9	2.2350	16	17	2.1547
17	9	9	2.2350	17	18	2.1547
18	10	10	2.2350	13	13	2.1547
19	10	10	2.2350	14	14	2.1547
20	10	10	2.2350	15	15	2.1547
21	10	10	2.2350	16	16	2.1547
22	10	10	2.2350	17	17	2.1547
23	11	11	2.2350	15	15	2.1547
24	11	11	2.2350	16	16	2.1547
25	11	11	2.2350	17	17	2.1547

Table A.2 Inflexible demand charging of 25 EV types.

Type	Total energy requirement (kWh)	Fixed charging rate (kW)	Fixed charging time (h)
1	4.3897	3	11-12
2	4.3897	3	12-13
3	4.3897	3	13-14
4	4.3897	3	14-15
5	4.3897	3	15-16
6	4.3897	3	16-17
7	4.3897	3	17-18
8	4.3897	3	18-19
9	4.3897	3	19-20
10	4.3897	3	12-13
11	4.3897	3	13-14
12	4.3897	3	14-15
13	4.3897	3	15-16
14	4.3897	3	16-17
15	4.3897	3	17-18
16	6.5445	3	18-20
17	6.5445	3	19-20
18	4.3897	3	14-15
19	4.3897	3	15-16
20	4.3897	3	16-17
21	4.3897	3	17-18
22	4.3897	3	18-19
23	4.3897	3	16-17
24	4.3897	3	17-18
25	4.3897	3	18-19