

# Bilevel Programming Approach to Demand Response Management with Day-ahead Tariff

András Kovács

**Abstract** This paper introduces a bilevel programming approach to electricity tariff optimization for the purpose of demand response management in smart grids. In the multi-follower Stackelberg game model, the leader is the profit-maximizing electricity retailer, who must set a time-of-use variable energy tariff in the grid. Followers correspond to groups of prosumers (simultaneous producers and consumers of electricity), who, in response to the observed tariff, schedule their controllable loads and determine the charging/discharging policy of their batteries to minimize their cost of electricity and to maximize their utility at the same time. A bilevel programming formulation of the problem is defined, and its fundamental properties are proven. The main contribution of the paper is the novel solution approach: primal-dual reformulation is proposed for converting the bilevel optimization problem into a single-level quadratically constrained quadratic program (QCQP), and a successive linear programming (SLP) algorithm is applied to solving it. It is demonstrated in computational experiments that the proposed approach outperforms typical earlier methods based on the KKT reformulation regarding both solution quality and computational efficiency on practically relevant problem sizes, while it also offers more flexible modeling capabilities.

**Keywords** Demand response management, smart grid, game theory, optimization, bilevel programming.

---

This research was supported by the GINOP grant No. 2.3.2-15-2016-00002, the NKFI grant 129178, and the János Bolyai Research Fellowship.

---

A. KOVÁCS, EPIC Center of Excellence in Production Informatics and Control, Institute for Computer Science and Control, Hungarian Academy of Sciences  
E-mail: akovacs@sztaki.hu

Received: date / Accepted: date

## 1 Introduction

A key to the stable operation of the future electricity grid is realizing efficient demand response management (DRM). With the increasing share of renewables in the energy mix, production is becoming less and less controllable. At the same time, electricity consumption is getting more controllable due to novel types of loads and storage (e.g., electric vehicles, home-level or small business energy management solutions) and various intelligent appliances at end consumers. As a result, a gradual shift from the traditional *supply follows demand* paradigm to a new, *demand follows supply* approach can already be observed. The critical success factor for efficient DRM is an appropriate electricity tariff that motivates consumers to schedule their loads and manage their batteries in such a way that it contributes to the stability of the grid.

This paper studies the problem of optimizing the electricity tariff offered by an electricity retailer to its customers in a game theoretical setting. A bilevel programming solution approach is introduced, where the retailer is the leader and the groups of end consumers act as multiple independent followers. The customers are modelled as *prosumers*, i.e., simultaneous producers and consumers of electricity, who look for the best tradeoff between maximizing their utility and minimizing their cost of electricity. An effective, and at the same time, computationally efficient solution method is proposed: the bilevel program is first transformed into an equivalent single-level optimization problem using a primal-dual reformulation, and then solved using a *successive linear programming* (SLP) algorithm.

The paper is organized as follows. After reviewing the related literature (Section 2), a formal definition of the tariff optimization problem is given (Section 3). The proposed solution approach is presented in detail in Section 4. The approach is illustrated on a small-scale ex-

ample and evaluated in thorough computational experiments in Section 5. Finally, conclusions are drawn and possible directions for future research are suggested.

## 2 Literature review

### 2.1 Game-theoretic models for DRM

Game-theoretic models for DRM have received significant attention recently [14]. A fundamental classification of these models differentiates between *real-time* vs. *day-ahead* approaches. *Real-time pricing* (RTP) models often, though not necessarily, focus on the present time instant only, and ignore the interdependence between the present energy tariff and past or future consumption. Accordingly, these models concentrate on load curtailment, but fail to capture deferrable loads appropriately. This limitation can be lifted by applying multi-period models. Still, most of the earlier contributions focus on the RTP scenario with a single time instant: a multi-leader, multi-follower Stackelberg game is defined for DRM among independent electricity providers and consumers in [23]; a closed-form analytical solution is derived, which can be obtained by a distributed algorithm. The management of consumer-to-grid systems is modelled as a Stackelberg game in [33], with a central power station acting as the leader, and consumers as multiple followers. Embedded into the Stackelberg game, the consumers play a generalized Nash game to establish their equilibrium strategies, and hence, to determine their response to the energy prices offered by the power station. A similar approach is applied to electric vehicle charging in [32]. Yu et al. [37] investigate DRM on three levels of hierarchy (the grid operator, multiple service providers, and the consumers) with real-time pricing, and propose a two-loop Stackelberg game model. The existence of a unique Stackelberg equilibrium is proven by exploiting the strictly convex sub-problems of the individual players, and an iterative distributed algorithm is proposed for reaching it. A Stackelberg approach is investigated for DRM under load uncertainty in [6]; again, an analytical solution could be derived.

Zugno et al. [39] study a Stackelberg game for RTP over multiple time periods with a profit-maximizing retailer and a single end consumer who looks for the best tradeoff between electricity cost and comfort in the heat management of a building. The problem is formulated as a bilevel program, and then converted to and solved as a single-level mixed-integer linear program (MILP) using the Karush-Kuhn-Tucker (KKT) conditions. The same paper shows that while real-time pricing is vastly

efficient for load shifting, it can cause excessive and imponderable payments for small consumers. Therefore, more predictable day-ahead pricing schemes are an attractive approach for households.

Despite this, the literature of *day-ahead* tariff optimization models is significantly scarcer. In [35], a Stackelberg model is proposed for energy pricing and dispatch in a multi-period day-ahead setting in two coupled stages. The first stage addresses price setting subject to demand response from consumers who minimize their energy cost and maximize their utility by scheduling their controllable loads. In the second stage, the retailer establishes the operating strategy for its storage unit and its energy contracts by solving a robust optimization problem considering uncertain market prices. The authors transform this problem into a single-level MILP by exploiting the KKT conditions and duality theory. In [36], a Stackelberg game is formulated and solved using an iterative heuristic approach. Two different games related to demand side management are studied in [31]; a Nash game between consumers equipped with batteries and a Stackelberg game between the utility provider and the consumers. A bilevel programming approach to the operation scheduling of a distribution network, with a cost-minimizing network operator as the leader and multiple profit-maximizing micro-grids as followers, is considered in [17]; again, KKT reformulation is applied to arrive at a single-level problem. A sophisticated Stackelberg game model is presented in [3] to capture the interplay of a retailer (leader) and various types of distributed energy resources, including generators and consumers with different types of controllable load as well (followers); again, the problem is converted to a single-level MILP using the KKT reformulation. A similar problem, with power flow constraints and a retailer who also oversees the operation of distributed generators and batteries, is studied in [24].

### 2.2 Related problems in energy management and DRM

The consumers' (followers') problem in the above Stackelberg games corresponds to an energy management problem for minimizing the cost and maximizing utility. Linear programming (LP) models limited to active power flow equations are commonly used in the literature for solving this problem [16]. More sophisticated, non-linear models allow capturing reactive power and voltage magnitudes as well [7, 29], or describe the behavior of the energy system components (battery, or heating, ventilation, and air conditioning in buildings) in a more realistic way [30].

Obviously, Stackelberg games and bilevel programming are not the only possible approaches to DRM problems. Alternative methods are often based on statistical models of the grid-level load response to the variation of energy prices, see, e.g. [9]. An iterative solution procedure that alternates between the optimization problems of the consumers (minimizing cost) and the grid (maximizing the load factor) is presented in [28] for smart building-to-grid systems, using a sophisticated thermal model of the buildings. In [22], the problem of dynamic pricing for DRM is formulated as a Markov decision process, and reinforcement learning is used to solving it. In [15], demand response is modeled by directly quantifying the delay-tolerant demand and its dependence on price by linear, potential, exponential and logarithmic load functions.

Game-theoretic approaches to different, but related problems in energy management include [1], where a Stackelberg approach is proposed for achieving a fair curtailment of renewable energy generation. A Stackelberg game model is investigated in [2] with a central production unit (leader) who sets the electricity price to maximize its profit subject to the response from an electricity service provider (follower) that will accept load curtailment and distributed generation bids from various micro-grids in view of the central producer's price. A supply-demand game is investigated in a smart grid setting in [38], with generators and loads acting as multiple followers, and a data center server as the virtual leader; a deep transfer Q-learning algorithm is applied for finding the equilibrium. The optimal operation of multi-carrier energy systems is modeled as bilevel optimization problem in [25]. The upper level problem of minimizing the total energy cost and the lower level problem of minimizing the operating and dissatisfaction costs are solved through an iterative procedure.

### 2.3 Mathematical methodology

An introduction to bilevel programming, including basic modelling and solution techniques, is given in [8, 11]. Approaches to transforming bilevel optimization problems into equivalent single-level models, including the optimal value or the KKT reformulation, are studied in [12]. A recent survey on bilevel programming for price setting problems is given in [21].

SLP has been applied in the smart grid community, e.g., to the planning of generators' investments and transmission network extensions [26], or to tackling non-linear phenomena in variants of the optimal power flow problem [5]. At the same time, to the best of our knowledge, this paper is the first to apply SLP to tariff optimization for demand response.

### 2.4 Positioning of the current contribution

This section surveyed the literature on game theoretical models to electricity tariff optimization for DRM and related problems. The algorithmic techniques applied to Stackelberg tariff optimization problems are summarized in a tabular format in Table 1. The survey shows that although some simpler formulations, all focusing on a single time period, adopt a solution that can be computed analytically in closed form, multi-period problems are computationally more challenging. This observation is also supported by a formal proof in [19], which states that multi-period models for DRM with controllable loads at the consumers are NP-hard. Accordingly, the vast majority of earlier contributions apply the KKT reformulation to arrive at a single-level MILP that can be solved using commercial solvers, however, at the price of considerable computational effort, and a number of papers mention that the solution approach is applicable mostly to small-scale problems [24]. Other approaches use customized heuristics for solving the problem.

With respect to the above state-of-the-art, the contribution of this paper is twofold. On the one hand, it defines a generic game-theoretic model for DRM that slightly extends the above discussed models (e.g., it captures both distributed battery storage and controllable loads characterized by a given utility function at the consumers), and whose key properties (e.g., necessary conditions for feasibility, computational complexity) can be formally proven. On the other hand, it proposes an efficient solution approach based on well-established mathematical programming techniques that first exploits duality for the followers' model to convert it into a single-level quadratically constrained quadratic program (QCQP), and subsequently, applies an SLP approach to solving it. It is shown in computational experiments that the proposed approach outperforms earlier KKT-based methods regarding both solution quality and computational effort for practically relevant problem sizes.

This paper is a substantially extended version of the earlier conference paper [18]. In addition to a refined model that captures the profit-maximizing behavior of the retailer, the extensions are related to the core contributions of the present paper, i.e., the formal proofs of the fundamental characteristics of the model and the thorough computational experimentation for assessing solution quality and computational efficiency.

**Table 1** Algorithmic techniques applied to for solving Stackelberg tariff optimization models in the literature.

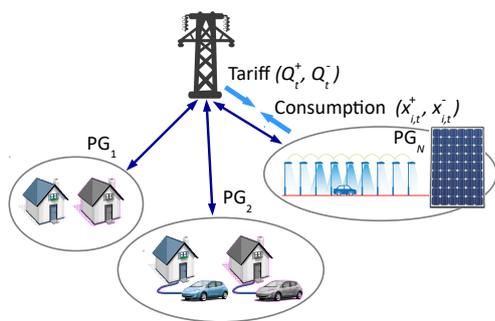
Paper	Game model	Time representation	Solution approach
[23]	Stackelberg game	Single period	Closed-form analytical solution*
[33, 32]	Generalized Nash embedded in Stackelberg	Single period	Iterative distributed algorithm using KKT & convex optimization
[37]	3-level Stackelberg game	Single period	Iterative distributed algorithm
[6]	Stackelberg game, stochastic	Independent periods**	Closed-form analytical solution
[39]	Stackelberg game	Multi-period	KKT for conversion to single level, MILP
[35]	Stackelberg game***	Multi-period	KKT for conversion to single level, MILP
[36]	Stackelberg game	Multi-period	Iterative algorithm
[31]	Stackelberg game****	Multi-period	Iterative algorithm
[17]	Stackelberg game	Multi-period	KKT for conversion to single level, MILP
[2]	Stackelberg game	Multi-period	KKT for conversion to single level, MILP
[24]	Stackelberg game	Multi-period	KKT for conversion to single level, MILP

\* A distributed algorithm for achieving the equilibrium is also presented.

\*\* The multi-period problem can be reduced to multiple single-period problems.

\*\*\* In a 2nd-stage problem, robust optimization is applied to implement the equilibrium under uncertainty.

\*\*\*\* A Nash game among consumers is also described in the paper.

**Fig. 1** System architecture with a retailer and multiple PGs.

### 3 Problem Definition

#### 3.1 System Architecture

This paper investigates DRM as an interaction among an electricity retailer and various prosumers, i.e., clients who can both produce and consume electricity, in a smart grid setting. In order to ensure the tractability of the problem over a large population, prosumers are classified into prosumer groups (PGs), where each PG consists of prosumers with similar electricity consumption and production profiles, as well as storage capabilities. The system architecture is displayed in Figure 1.

PGs are characterized by their uncontrollable production and consumption, controllable load requirements, as well as their storage capabilities. The uncontrollable production  $C_{i,t}^+$  and consumption  $C_{i,t}^-$  of  $PG_i$  is fixed and given in the input for each time period  $t = 1, \dots, T$ . In addition to that,  $PG_i$  needs to schedule a (potentially zero) controllable load of  $M_i$  over the time horizon, where a maximum of  $\bar{L}_{i,t}$  can be scheduled in each period  $t$ . It is noted that time windows can be defined for the controllable load by setting  $\bar{L}_{i,t} = 0$  for the appropriate periods  $t$ . The preferences of  $PG_i$  on the timing of the controllable load are encoded in utility

values  $U_{i,t}$ , where  $U_{i,t}$  captures (the monetary equivalent of) the utility of scheduling a unit of controllable load in time period  $t$ . Hence, if  $PG_i$  decides for a controllable load of  $L_{i,t}$  over time, then this incurs a utility of  $\sum_{t=1}^T U_{i,t} L_{i,t}$  for the PG. Similar models for controllable load are used frequently in the literature, see, e.g., [35].

PGs can further optimize their energy management by the appropriate charging and discharging of their battery storage. The battery is characterized by its capacity  $\bar{B}_i$ , the maximum charge and discharge rates  $R_i^+$  and  $R_i^-$ , the initial battery state-of-charge (SoC)  $b_{i,0}$ , and its cycle efficiency  $\eta_i$ . In order to safeguard from unexpected power outages, the prosumer wishes to retain a given, time-varying minimum state-of-charge  $\underline{B}_{i,t}$  in the battery.

Each individual PG schedules its controllable load  $L_{i,t}$  and determines its battery state-of-charge  $b_{i,t}$  over time to optimize its own objective, composed of maximizing the utility and minimizing the electricity cost w.r.t. the energy tariff set by the retailer. This PG model is generic enough to capture the behavior of diverse types of prosumers, ranging from households or offices with uncontrollable consumption only (and therefore, unresponsive to the energy tariff), via prosumers equipped with renewable energy generation and/or storage devices, owners of electric vehicles, to complex micro-grid systems.<sup>1</sup>

The retailer employs the same time-of-use electricity tariff for all prosumers. The tariff is specified in

<sup>1</sup> It should be noted that various alternative approaches to modeling prosumer behavior have been subject to extensive research recently. Questions of special interest include addressing individual prosumers or organizing them into PGs, as well as using deterministic or probabilistic models. A richer, probabilistic approach to characterizing the responsiveness of prosumers to the variation of the electricity tariff is presented, e.g., in [34], together with a review of the recent literature on the benefits and drawbacks of different approaches.

the form of day-ahead electricity purchase prices  $Q_t^+$  and feed-in prices  $Q_t^-$  for periods  $t = 1, \dots, T$ . It is assumed that the tariff is regulated by an a priori agreement between the retailer and the prosumers, which defines minimum, maximum, and max. average electricity prices in the form of  $0 < \underline{Q} \leq Q_t^- \leq Q_t^+ \leq \bar{Q}$  and  $\frac{1}{T} \sum_{t=1}^T Q_t^+ \leq \tilde{Q}$ . Such an agreement is necessary to prevent the profit maximizing retailer from increasing purchase prices without any limit, see, e.g., [39].

The focus of this paper is on the problem faced by the retailer, who has to cover the (potentially negative) net consumption of the ensemble of all prosumers from electricity purchased (or sold) on the wholesale market. This paper assumes a time-variant dual pricing scheme on the wholesale market, given in the form of purchase prices  $P_t^+$  and feed-in prices  $P_t^-$ . It is noted that the same model can be naturally applied to markets with uniform pricing (purchase prices equal to selling prices) by letting  $P_t^+ = P_t^-$ . It is assumed that the retailer appears as a price-taker on the market, without any ability to influence the market prices.

**Table 2** Notation used in the paper.

Dimensions	
$T$	Number of time periods
$N$	Number of prosumer groups (PGs)
Parameters for PG <sub><i>i</i></sub>	
$C_{i,t}^+$	Uncontrollable production [kWh]
$C_{i,t}^-$	Uncontrollable consumption [kWh]
$M_i$	Total controllable load during the horizon [kWh]
$\bar{L}_{i,t}$	Maximum controllable load scheduled [kWh]
$U_{i,t}$	Utility of controllable load scheduled [\$/kWh]
$R_i^+$	Maximum battery charge rate [kWh/period]
$R_i^-$	Maximum battery discharge rate [kWh/period]
$\eta_i$	Cycle efficiency of the battery [%]
$\bar{B}_i$	Battery capacity [kWh]
$\underline{B}_{i,t}$	Minimum battery SoC [kWh]
$b_{i,0}$	Initial battery SoC [kWh]
Parameters for the retailer	
$\bar{Q}$	Max. average electricity price for prosumers [\$/kWh]
$\underline{Q}$	Maximum electricity price for prosumers [\$/kWh]
$\tilde{Q}$	Minimum electricity price for prosumers [\$/kWh]
$\bar{P}^+$	Purchase price of electricity on wholesale market [\$/kWh]
$\bar{P}^-$	Selling price of electricity on wholesale market [\$/kWh]
Decision variables for PG <sub><i>i</i></sub>	
$x_{i,t}^+$	Electricity purchased [kWh]
$x_{i,t}^-$	Electricity fed into the grid [kWh]
$L_{i,t}$	Controllable load [kWh]
$r_{i,t}^+$	Electricity charged into battery [kWh]
$r_{i,t}^-$	Electricity discharged from battery [kWh]
$b_{i,t}$	Battery SoC at the end of period $t$ [kWh]
$\varphi_{i,t}^k$	Dual variables (see definition at primal constraints)
Decision variables for the retailer	
$Q_t^+$	Electricity purchase price offered to PGs [\$/kWh]
$Q_t^-$	Electricity feed-in price offered to PGs [\$/kWh]
$y_t^+$	Electricity purchased on the wholesale market [kWh]
$y_t^-$	Electricity sold on the wholesale market [kWh]
Objective values	
$f$	Objective value of the retailer [\\$]
$g_i$	Objective value of PG $i$ [\\$]

The retailer, by offering an appropriate electricity tariff to its prosumers, can initiate a demand response program that motivates the prosumers to purchase electricity in valley periods, when an ample amount of cheap energy is available on the market, and sell their surplus energy in peak periods. In this way, the retailer can contribute to the stability of the grid and maximize its profit at the same time. In this paper, the maximization of the retailer's profit is addressed. The applied notation is summarized in Table 2.

### 3.2 Stackelberg Game Model and Basic Characteristics

The following communication protocol is implemented among the various stakeholders: the retailer first announces the day-ahead electricity tariff to all prosumers. The prosumers observe this tariff and optimize their consumption profile, i.e., the amount of electricity purchased from or fed into the grid over time. Then, the parties implement their actions as planned. It is assumed that the retailer is aware of the decision model and the parameters of the PGs. This leads to a *Stackelberg game* with the retailer as the *leader* and the PGs as *multiple followers*. The so-called *optimistic* assumption is adopted, i.e., if a follower has more than one optimal solutions according to its own objective, then it chooses its optimal solution that is the *most favorable* for the leader. The following additional assumptions are made, which guarantee the feasibility of the problem:

$$\sum_{t=1}^T \bar{L}_{i,t} \geq M_i \quad \forall i \quad (1)$$

$$\underline{B}_{i,t} \leq \bar{B}_i \quad \forall i, t \quad (2)$$

$$b_{i,0} \leq \bar{B}_i \quad \forall i \quad (3)$$

$$b_{i,0} + tR_i^+ \leq \underline{B}_{i,t} \quad \forall i, t \quad (4)$$

These assumptions require that bounds on the controllable load allow scheduling the required power over the horizon (1), the bounds on the battery state-of-charge are consistent (2), the initial battery charge satisfies these bounds (3), and finally, that the charging rate of the battery allows satisfying the lower bounds on the state-of-charge (4).

**Lemma 1 (Existence of a solution)** *If assumptions (1)-(4) hold, then the followers' problem is feasible for any electricity tariff  $Q_t^+$  and  $Q_t^-$  set by the leader.*

**Proof:** Setting the battery state-of-charge to the required minimum, i.e.,  $b_{i,t} = \max(b_{i,t-1}, \underline{B}_{i,t})$ , and scheduling the controllable as early as the bounds allow, i.e.,

$L_{i,t} = \min \left( \bar{L}_{i,t}, M_i - \sum_{u=1}^{u=t-1} L_{i,u} \right)$  results in a feasible solution for each follower  $i$ .  $\square$

**Lemma 2 (Independence of followers’ problems)**

*The optimal demand response of an individual PG to a given energy tariff is independent of the response of other PGs.*

**Proof:** The objectives of the individual PG, i.e., its energy cost and utility, depend solely on the energy tariff and the consumption profile of the given PG. Moreover, the feasibility of a consumption profile is also independent of other PGs’ response, since the amount of electricity that can be purchased or sold on the market by the retailer to maintain the grid-level balance is unbounded.  $\square$

It is emphasized that different PG’s problems are still interconnected through the retailer’s problem, but for any fixed decision of the retailer on the tariff, the PGs can optimize their behavior without considering the problems faced by fellow PGs. Hence, the problem can be modeled as a Stackelberg game with a single leader (the retailer) and multiple independent followers (the PGs). It is noted that when the optimal response of a follower is not unique, then the response induced by the optimistic assumption (from the set of all optimal responses) can be dependent on other PGs’ response.

**Lemma 3 (Computational complexity)** *The above defined bilevel energy tariff optimization problem is  $\mathcal{NP}$ -hard.*

**Proof:** The *Simple Multi-period Energy Tariff Optimization Problem* (SMETOP) has been introduced as a minimal bilevel optimization model of energy tariff optimization for DRM, and it has been proven to be  $\mathcal{NP}$ -hard in [19]. The problem investigated in this paper generalizes SMETOP in the sense that, in addition to all features captured by SMETOP, it also handles batteries and uncontrollable energy production and consumption at the PGs, as well as bi-directional grid connections. This implies that the currently investigated, generalized problem is  $\mathcal{NP}$ -hard, too.  $\square$

## 4 Solution Approach

### 4.1 Overview

This section presents a bilevel programming formulation of the above Stackelberg game model, and proposes an efficient solution approach for that formulation. First, the models of an individual follower and the

leader are formally defined. Then, the bilevel programming model received as a combination of the two parties’ problems is reformulated into a single-level QCQP, which is, in turn, solved using an SLP algorithm.

### 4.2 Prosumer Groups’ (Followers’) Problem

The decision problem faced by an individual PG <sub>$i$</sub>  (follower) is a parametric optimization problem, whose parameters encode the electricity tariff determined by the retailer (decision variables  $Q_t^+$  and  $Q_t^-$  controlled by the leader). The problem can be captured by the following linear program (LP), where symbols  $(\varphi_{i,t}^k)$  on the r.h.s. of the constraints denote the dual variables associated with the given constraint:

Minimize

$$g_i(Q^+, Q^-) = \sum_{t=1}^T (Q_t^+ x_{i,t}^+ - Q_t^- x_{i,t}^- - U_{i,t} L_{i,t}) \quad (5)$$

subject to

$$C_{i,t}^+ - C_{i,t}^- + x_{i,t}^+ - x_{i,t}^- - L_{i,t} = r_{i,t}^+ - r_{i,t}^- \quad \forall t \quad [\varphi_{i,t}^6] \quad (6)$$

$$\eta_i r_{i,t}^+ - r_{i,t}^- = b_{i,t} - b_{i,t-1} \quad \forall t \quad [\varphi_{i,t}^7] \quad (7)$$

$$\sum_{t=1}^T L_{i,t} = M_i \quad [\varphi_i^8] \quad (8)$$

$$L_{i,t} \leq \bar{L}_{i,t} \quad \forall t \quad [\varphi_{i,t}^9] \quad (9)$$

$$\underline{B}_{i,t} \leq b_{i,t} \quad \forall t \quad [\varphi_{i,t}^{10}] \quad (10)$$

$$b_{i,t} \leq \bar{B}_i \quad \forall t \quad [\varphi_{i,t}^{11}] \quad (11)$$

$$r_{i,t}^+ \leq R_i^+ \quad \forall t \quad [\varphi_{i,t}^{12}] \quad (12)$$

$$r_{i,t}^- \leq R_i^- \quad \forall t \quad [\varphi_{i,t}^{13}] \quad (13)$$

$$0 \leq x_{i,t}^+, x_{i,t}^-, r_{i,t}^+, r_{i,t}^-, L_{i,t} \quad \forall t \quad (14)$$

The follower’s objective (5) is comprised of the total cost of energy, i.e., the cost of energy purchased minus the income from feeding energy into the grid, and the PG’s utility achieved by the timing of the controllable load. Constraint (6) encodes that the energy balance at the PG is maintained. Equation (7) computes the battery state-of-charge based on the charge and discharge

rates. Constraints (8) and (9) ensure that the amount and the timing of the controllable load satisfies the requirements. Finally, inequalities (10)-(14) define the allowed range of the battery state-of-charge, the charge and discharge rates, as well as the electricity purchase and feed-in rates at the PG.

It is noted that all constraints in the followers' model are linear, whereas the objective contains the leader's variables as multipliers, making it a bilinear (quadratic) expression. The models of different followers are linked only via the leader's decision problem.

#### 4.3 Retailer's (Leader's) Problem

The optimization problem faced by the retailer can be formulated as a bilevel program that contains the PGs' problem as a nested sub-problem. This nested sub-problem, encoded as a constraint in the model, expresses that a part of the variables (namely, decision variables  $x_{i,t}^+$  and  $x_{i,t}^-$ , corresponding to the amount of electricity purchased from and fed into the grid) are controlled by the followers, according to their known decision model:

Maximize

$$f = \sum_{t=1}^T \left( \sum_{i=1}^N (Q_t^+ x_{i,t}^+ - Q_t^- x_{i,t}^-) - P_t^+ y_t^+ + P_t^- y_t^- \right) \quad (15)$$

subject to

$$y_t^+ - y_t^- = \sum_{i=1}^N (x_{i,t}^+ - x_{i,t}^-) \quad \forall t \quad (16)$$

$$Q \leq Q_t^- \leq Q_t^+ \leq \bar{Q} \quad \forall t \quad (17)$$

$$\frac{1}{T} \sum_{t=1}^T Q_t^+ \leq \tilde{Q} \quad (18)$$

$$\begin{pmatrix} x_{i,t}^+ \\ x_{i,t}^- \end{pmatrix} \in \arg \min \{ g_i(Q^+, Q^-) \mid (6) - (14) \} \quad \forall i \quad (19)$$

The leader's objective (15) is maximizing its profit, calculated as its revenue from the prosumers, minus the cost of electricity purchased on the market, plus the income from the electricity sold on the market. Equation (16) encodes the grid-level energy balance. Inequalities (17) and (18) define the valid range of the energy tariff variables. Finally, constraint (19) states that the prosumers' electricity purchase and feed-in values are determined by them using the above optimization model.

#### 4.4 Single-level QCQP Reformulation

When the complexity of a bilevel optimization problem does not allow developing an analytical solution, which is apparently the case above, the two candidate solution approaches are the application of (meta-)heuristics directly to the bilevel problem, or reformulation to a single-level problem. The considerable benefit of the latter technique is that it allows the application of theoretically well-founded, potentially even exact, mathematical programming approaches to solving the problem. For this reason, this paper adopts the reformulation approach and looks for a transformation of the bilevel problem (15)-(19) into a single-level mathematical program. The key to achieving this is modeling the followers' optimality condition (19). By exploiting duality for the followers' LP model<sup>2</sup>, a primal-dual reformulation of the followers' problem can be applied: the optimality condition (19) is translated into the conjunction of the followers' primal constraints (6)-(14), dual constraints, and an equality constraint between the primal and the dual objectives. By duality, the ensemble of these constraints is satisfied if and only if the given instantiation of the variables is an optimal solution for the follower.

The complete single-level reformulated problem is shown below. It consist of the leader's objective (15), the leader's constraints (16)-(18), the followers' primal constraints (6)-(14), an equality relation between the followers' primal and dual objectives (20), as well as the followers' dual constraint corresponding to the primal variables for the battery charge rate  $r_{i,t}^+$  (21), discharge rate  $r_{i,t}^-$  (22), state of charge  $b_{i,t}$  for  $t < T$  (23) and  $b_{i,T}$  (24), electricity purchase  $x_{i,t}^+$  (25), electricity feed-in  $x_{i,t}^-$  (26), and controllable load  $L_{i,t}$  (27).

Maximize (15)

subject to

$$(6) - (14), (16) - (18)$$

$$\sum_{t=1}^T (Q_t^- x_{i,t}^- - Q_t^+ x_{i,t}^+ + U_{i,t} L_{i,t})$$

$$= \sum_{t=1}^T \left( (C_{i,t}^- - C_{i,t}^+) \varphi_{i,t}^6 + \bar{L}_{i,t} \varphi_{i,t}^9 + R_i^+ \varphi_{i,t}^{12} + R_i^- \varphi_{i,t}^{12} - \underline{B}_{i,t} \varphi_{i,t}^{10} + \bar{B}_i \varphi_{i,t}^{11} \right) \quad \forall i \quad (20)$$

<sup>2</sup> It is noted that the followers' LP model (5)-(14) contains the bilinear term  $Q_t^+ x_{i,t}^+ - Q_t^- x_{i,t}^-$  in the expression of  $g_i(Q^+, Q^-)$ , i.e., a multiplication of the leader's and the followers' variables. With this, the model is still linear in the followers' variables, and hence, LP duality can be exploited for reformulating it.

$$-\varphi_{i,t}^6 + \eta_i \varphi_{i,t}^7 + \varphi_{i,t}^{12} \geq 0 \quad \forall i, t \quad (21)$$

$$\varphi_{i,t}^6 - \varphi_{i,t}^7 + \varphi_{i,t}^{13} \geq 0 \quad \forall i, t \quad (22)$$

$$-\varphi_{i,t}^7 + \varphi_{i,t+1}^7 - \varphi_{i,t}^{10} + \varphi_{i,t}^{11} \geq 0 \quad \forall i, t < T \quad (23)$$

$$-\varphi_{i,T}^7 - \varphi_{i,T}^{10} + \varphi_{i,T}^{11} \geq 0 \quad \forall i \quad (24)$$

$$\varphi_{i,t}^6 \geq -Q_t^+ \quad \forall i, t \quad (25)$$

$$-\varphi_{i,t}^6 \geq Q_t^- \quad \forall i, t \quad (26)$$

$$-\varphi_{i,t}^6 + \varphi_{i,t}^8 + \varphi_{i,t}^9 \geq U_{i,t} \quad \forall i, t \quad (27)$$

$$\varphi_{i,t}^9, \varphi_{i,t}^{10}, \varphi_{i,t}^{11}, \varphi_{i,t}^{12}, \varphi_{i,t}^{13} \geq 0 \quad \forall i, t \quad (28)$$

The primal-dual reformulation particularly suits the problem in scope, since the only occurrence of the leader's variables,  $Q_t^+$  and  $Q_t^-$ , in the followers' problem is in their primal objective, and consequently, on the right hand side of the dual constraints. As a result, the only non-linear term in the single-level reformulation is the payment from the PGs to the retailer, contained both in the leader's objective (15) and in the followers' optimality constraint (20), which is a bilinear expression containing the multiplication of the followers' and the leader's variables. All other constraints are linear.

#### 4.5 SLP Solution Method

Since the above QCQP is non-convex, no efficient exact algorithm can be expected for solving it, and accordingly, (meta-)heuristic approaches are of interest. Therefore, we propose an SLP heuristic solution approach, which is known to show good convergence properties especially on problems where most of the constraints are linear, which is the case above. SLP solves non-linear problems by iteratively constructing local LP approximations of the original problem, and solving each approximation using standard LP techniques [4, 27]. The algorithm departs from an initial solution  $X_0$ , and in each iteration  $k$ , it builds a local linearization of the original problem around  $X_k$ , denoted as  $LP_k$ . Then, the optimal solution of  $LP_k$  is sought subject to a given step bound,  $-s \leq X - X_k \leq s$ . If the optimal LP solution is feasible with a given tolerance, then it is accepted as the next solution  $X_{k+1}$  (and possibly  $s$  is increased); otherwise  $X_{k+1} = X_k$  and  $s$  is decreased.

The above SLP algorithm converges to a locally optimal solution of the QCQP, which is potentially different from the global optimum. In order to reduce the risk of getting stuck in a local optimum, the SLP algorithm is embedded into a randomized restart procedure. It executes multiple SLP runs, using a random perturbation of the previous best solution as an initial solution in each run (or  $Q_t^- = Q_t^+ = \underline{Q}$  in the first run). The implementation reported in this paper is based on the SLP package of Fico Xpress 7.8, using its default SLP

algorithm, with the number of SLP runs set to 10 in all computational experiments.

#### 4.6 Discussion on Possible Extensions

While the above presented bilevel model captures the most important generic features of prosumers (production and consumption, controllable load, battery storage), it can be extended and refined in many different ways. The most relevant directions include the extension of the prosumer model with features for specific types of equipment that induce elastic load (e.g., HVAC in buildings, or refined battery storage models capturing state-dependent charging properties and losses [30]), as well as the extension of the retailer model with generation or energy storage. The proposed solution method is directly applicable to the extended models as long as the prosumer model remains linear. The proposed reformulation still applies with binary variables in the retailer model (e.g., due to switchable generators), and commercial solvers offer algorithms for tackling the resulting mixed-integer QCQP, though the computational efficiency of the approach needs to be verified for the given application.

Below, we review two minor refinements of the baseline bilevel model (15)-(19) that fix specific issues that might be undesirable in some application scenarios. First, the baseline model may trigger inappropriate end-of-horizon effects, namely, followers selling all the energy stored in the batteries to maximize their revenue. This can be avoided by subtracting a term that values the energy stored in the batteries at the end of the planning horizon from the followers objective (5) as follows:

$$\frac{Q_T^+ + Q_T^-}{2} b_{i,T} \quad (29)$$

Another example of a possible requirement that is not captured readily by the above model is that, among the different optimal solutions that maximize the retailer's profit, a solution with a smooth electricity purchase and/or sale over time is preferred. Unwanted oscillation of the energy purchased or sold on the wholesale market can be smoothed out by adding the following term to the retailer's objective:

$$-\varepsilon \sum_{t=1}^T (y_t^+ - y_t^-)^2 \quad (30)$$

This quadratic term measures the squared deviation of the energy traded over time (with a constant bias). Accordingly, adding it to the retailer's objective with a small multiplier  $\varepsilon$  smooths unnecessary oscillations without affecting the payoffs of the players.

#### 4.7 Discussion on the KKT Reformulation

As an alternative to the proposed solution approach, KKT reformulation and linearization can be applied to convert the proposed bilevel model (15)-(19) into a single-level MILP. This approach is often considered to be the default choice for transforming bilevel problems into single-level ones. Moreover, the resulting MILP, in theory, can be solved to exact optimality by commercial solvers.

Converting the bilevel model into a single-level MILP requires linearizing the KKT complementary slackness conditions using *big-M* constraints over additional binary variables, as well as linearizing the quadratic term in the objective by expressing and substituting it from (20). However, as it will be shown in the computational experiments, this approach is extremely challenging computationally due to the high number of binary variables and big-M constraints. In particular, linearizing the complementary slackness conditions requires introducing ca.  $22 \cdot NT$  auxiliary binary variables into the model (one for each primal and dual variable, resulting in over 20 000 additional binary variables for  $N = 20$  and  $T = 48$ ). Moreover, the corresponding big-M constraints are typically difficult to solve due to their weak LP relaxations. For further details on the KKT reformulation, the interested reader is referred to [8, 12]. Finally, even minor modifications in the bilevel model can hinder the linearization of the KKT reformulation, as is the case with the terms for the valuation of the remaining charge (29) or for smoothing (30).

In this paper, we use the KKT reformulation and the exact MILP solution approach in computational experiments to assess the quality of the solutions found by the proposed SLP solution approach on small-size problems.

## 5 Experimental Evaluation

### 5.1 Illustrative Example

In this section, the proposed approach to DRM is demonstrated on a small-scale illustrative example, with three PGs and a one-day horizon (from 08:00 AM, using hourly time units). The PGs correspond to different types of consumers as follows:

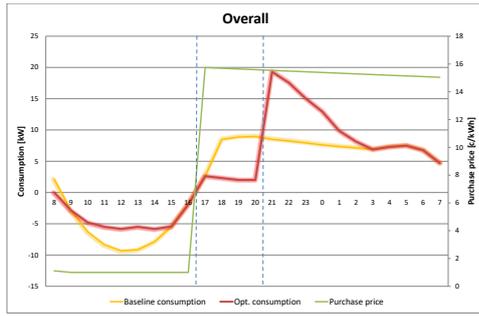
- $PG_1$  represents an intelligent energy-positive street lighting micro-grid system called E+grid [10, 20]. Since the lighting system is controlled according to local traffic and environmental conditions, as captured by motion sensors and a local weather station, its consumption varies dynamically over time.

The grid-connected system is also equipped with PV power generation and battery storage, which enables it to perform active energy management using an optimization approach that corresponds to the PG model adopted in this paper. Real-life data originates from a physical prototype with 191 luminaries and 151.2 m<sup>2</sup> of active PV surface area, and reflects the operation of the system on a sunny day in October. The E+grid PG is a net producer (up to 15 kW) during the day, and a net consumer (up to 3.5 kW) during the night.

- $PG_2$  comprises owners of plug-in electric vehicles. Data used in the example corresponds to three Nissan Leaf EVs, with a 24kWh battery pack in each vehicle, which has to be charged from a 50% state to 100% during the night. Individual vehicles are connected to the grid between 17:00–20:00 and disconnected between 6:00–8:00 in the morning. With the vehicle-to-grid (V2G) option ignored, this can be modeled as a controllable load of 36kWh. It is assumed that the owners have a slight preference for charging the EVs as early as possible, which is captured by utility values  $U_{2,t}$  linearly decreasing over time.
- Finally,  $PG_3$  contains households with uncontrollable consumption only. This case study uses the data of 15 average Hungarian homes, with a peak consumption of 5.7 kW during the day, and a minimum consumption of 3.8 kW during the night. Since this PG has no controllable load or battery storage, it cannot participate actively in DRM, and its consumption appears only as a time-varying bias in the grid-level consumption.

The retailer aims to maximize its profit by offering an appropriate time-of-use electricity tariff to the PGs, respecting the a priori contract that sets  $\underline{Q} = 1$  c/kWh,  $\bar{Q} = 100$  c/kWh, and  $\tilde{Q} = 10$  c/kWh. For the sake of simplicity, the wholesale market prices are assumed to vary in two steps: 12 c/kWh during the day (between 8:00 and 21:00) and 6 c/kWh during the night. The feed-in price on the wholesale market is a constant 3 c/kWh.

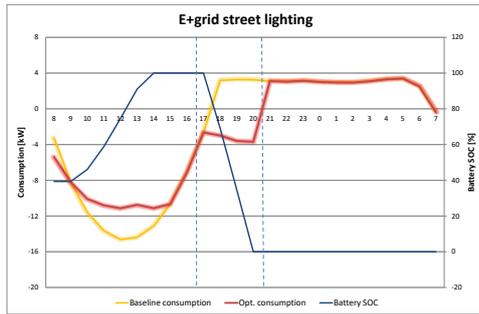
The system-level optimum for this example is determined by the following characteristics: the overall grid is a net producer until 17:00 due to PV generation in the E+grid micro-grid, whereas it is a net consumer afterwards. In order to avoid losses stemming from dual pricing on the wholesale market, the retailer should motivate the PGs to anticipate load and charge batteries before 17:00. On the contrary, in the period after 17:00, it should encourage PGs to defer their load from the peak period lasting until 21:00 to the valley period afterwards.



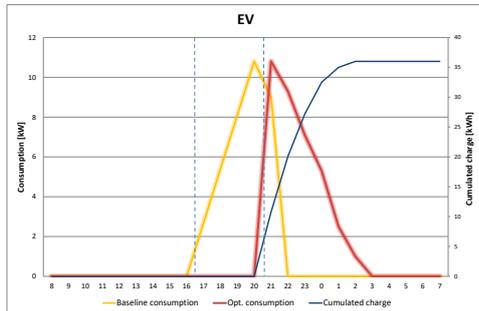
**Fig. 2** Solution with optimized tariff: energy purchase price and overall consumption over time.



**Fig. 5** Solution with optimized tariff: consumption of the households PG over time.



**Fig. 3** Solution with optimized tariff: consumption and battery state of the E+grid lighting system PG over time.



**Fig. 4** Solution with optimized tariff: consumption and cumulated load of the electronic vehicle PG over time.

This sample instance was solved using the proposed approach, applying formula (30) to eliminate the oscillations of the energy flow that are visually disturbing, but otherwise does not affect the payoffs of the players. The results displayed in Figures 2-5 show that the proposed approach could indeed reach the above described system-level optimum. The diagrams compare the optimized consumption profile (red curve) to the baseline consumption (yellow curve) for the overall grid and for the individual PGs, where the baseline consumption is computed by scheduling the controllable loads to maximize utility (ignoring the electricity tariff) and not using the batteries. The characteristic time periods are separated by dashed lines at 17:00 and 21:00. Finally, the optimized purchase tariff is also shown in the diagram of the overall grid: constant low prices (1 c/kWh)

are applied while the system is a net producer, until 17:00, whereas high, slightly decreasing prices are used afterwards (15.74 c/kWh at 18:00, decreasing by 0.05 c/kWh per hour).

On the level of individual PGs, the applied tariff motivated the E+grid PG to charge its battery while it is a net producer, to reach a fully charged state during 14-17:00, and to gradually discharge the battery in the rest of the peak period between 17:00 and 21:00. The controllable load of the EV PG was fully deferred to the valley period after 21:00. In that period, the slight decrease of the purchase prices over time compensates the PG for its linearly decreasing utility function, and therefore, an arbitrary scheduling of the controllable load became optimal for this PG. There were no controllable variables for the households PG. This tariff and consumption profile is globally optimal for the retailer, since no further load can be moved outside the peak period between 17:00 and 21:00.

The implementation of the proposed solution approach in Fico Xpress 7.8, with 10 SLP runs within the randomized restart heuristic, could solve the above problem instance in 9.8 seconds on a computer with Intel i7 2.70GHz CPU and 16GB RAM.

## 5.2 Computational Experiments

The evaluation of the proposed approach in computational experiments focused on two questions: (1) the computational effort required by the proposed SLP solution approach, and (2) the quality of the solutions found. Large problem instances were generated by the multiplication and random perturbation of the data used in the above illustrative example. Table 3 displays the average computation time in seconds over 10 instances for different combinations of  $N$  (number of PGs) and  $T$  (number of time periods), achieved with the proposed model (15)-(19) and algorithm using 10 SLP runs. The results show that computation time increases moderately with problem size, and practically

**Table 3** Average computation time in seconds by problem size for the proposed SLP solution approach.

	$T = 12$	$T = 24$	$T = 36$	$T = 48$
$N = 5$	0.55	1.46	2.22	16.04
$N = 10$	1.10	73.38	57.97	228.37
$N = 15$	1.67	155.74	51.59	283.56
$N = 20$	3.63	257.31	98.82	345.06

**Table 4** Comparison of the exact KKT and the proposed SLP solution approaches. Negative gaps arise where SLP outperforms KKT within the 600 s time limit.

Size $N \cdot T$	KKT		SLP			
	Opt	Time (s)	Gap			Time (s)
			min.	avg.	max.	
3 · 12	100%	0.51	0.00%	0.00%	0.00%	0.43
3 · 24	100%	4.43	0.00%	0.09%	0.98%	1.11
3 · 36	100%	55.00	0.00%	0.01%	0.06%	1.29
3 · 48	90%	240.64	-19.16%	-1.93%	0.05%	2.24
5 · 12	60%	304.01	-0.04%	0.00%	0.02%	0.55
5 · 24	0%	600.00	-6.99%	-2.11%	0.83%	1.46
5 · 36	0%	600.00	-40.92%	-8.30%	0.48%	2.22
5 · 48	0%	600.00	-28.13%	-5.93%	1.51%	16.04

relevant problem sizes, e.g., with  $N = 20$  and  $T = 48$ , were tractable in a reasonable amount of time. In applications where a different tradeoff between solution quality and computation time is looked for, the algorithm can be tuned, e.g., by modifying the number of SLP runs.

In order to evaluate the quality of the solutions found by the proposed approach, they were compared to the exact optimal solutions of the MILP model received by applying KKT reformulation and linearization as discussed in Section 4.7. The results of the comparison are displayed in Table 4, which displays aggregated results over 10 instances for each problem size. Column  $KKT/Opt$  contains the ratio of instances that could be solved to proven optimality using KKT, and  $KKT/Time$  shows the average computation time required for this. The branch and bound search was aborted when the time limit of 600 s was hit, and the best integer solution found was recorded. Columns  $SLP/Min. gap$ ,  $SLP/Avg. gap$ , and  $SLP/Max. gap$  display the minimum, the average, and the maximum gap between the SLP and the KKT solutions for the given problem size. Finally, column  $SLP/Time$  contains the average computation time for the SLP solution approach. The results show that the smallest instances, with  $N = 3$ , could be solved to proven optimality using KKT (with a single exception). Although SLP is not an exact solution approach, in practice it also builds close-to-optimal solutions with an average gap of only 0.01-0.1%. For somewhat larger instances,  $N = 5$ , where KKT failed to find the optimal solution, SLP often constructed significantly better solutions, as indicated by negative gap values. Namely, SLP found up to 40% better solutions

than KKT with one or two orders of magnitude lower computation times.

Hence, it can be concluded that although KKT is an exact solution approach in theory, in practice its applicability is limited to small problems, e.g., with  $N = 3$ . In contrast, the proposed primal-dual reformulation coupled with SLP scales much more favorably, and it computes high-quality solutions efficiently even for practically relevant problem sizes.

## 6 Conclusions

This paper introduced a bilevel programming approach to energy tariff optimization for DRM in smart grids. In the Stackelberg game model, the leader is a profit maximizing retailer, who sets the energy tariff offered to its prosumers and purchases electricity for them from the wholesale market. The prosumers, who act as multiple independent followers, optimize their controllable load and their battery charging schedule to maximize their utility and minimize their cost of energy. A novel solution approach was introduced, which exploits the primal-dual reformulation of the followers' problem to arrive at a single-level QCQP equivalent of the bilevel problem. It has been shown that the resulting QCQP can be solved efficiently using an SLP algorithm. In particular, it was illustrated in computational experiments that the proposed approach outperforms the technique based on the KKT reformulation, which is the dominant approach for solving similar problems in the literature. Hence, the main contributions of the paper are a bilevel programming formulation of the tariff optimization problem, formal proofs of some basic properties, and the application of novel and efficient mathematical programming techniques to solve this problem.

The proposed model can be trivially extended to some more complex problems, e.g., with various types of controllable loads and storage devices for each PG, or switchable generators and energy storage at the retailer. A more important and challenging direction for future research is the investigation of richer, non-linear prosumer models that can capture more realistically, e.g., thermal processes of HVAC in buildings or charging properties of batteries. The extension to a stochastic variant, accounting for uncertainties in consumption, production, and spot market prices is also of interest.

Finally, it must be observed that while Stackelberg game models are becoming ubiquitous in the literature of DRM, a critical pre-condition of their practical applicability is that the leader should be able to identify the decision models and parameters the followers. This is a challenging problem in application scenarios characterized by information asymmetry. A promising solution

approach can be the application of inverse optimization, analogously to a case in inventory control [13]; given historical pairs of a follower's input (i.e., energy tariff) and response (consumption), the inverse optimization approach looks for parameters that ensure that each response is optimal for the corresponding input.

## References

- [1] M. Andoni, V. Robu, W-G. Früh, and D. Flynn. Game-theoretic modeling of curtailment rules and network investments with distributed generation. *Applied Energy*, 201:174–187, 2017.
- [2] G. E. Asimakopoulou, A. L. Dimeas, and N. D. Hatziargyriou. Leader-follower strategies for energy management of multi-microgrids. *IEEE Trans. Smart Grid*, 4(4):1909–1916, 2013.
- [3] G. E. Asimakopoulou, A. G. Vlachos, and N. D. Hatziargyriou. Hierarchical decision making for aggregated energy management of distributed resources. *IEEE Trans. Power Systems*, 30(6):3255–3264, 2015.
- [4] R. H. Byrd, N. I. M. Gould, J. Nocedal, and R. A. Waltz. An algorithm for nonlinear optimization using linear programming and equality constrained subproblems. *Mathematical Programming, Ser. B.*, 100(1):27–48, 2003.
- [5] A. Castillo, P. Lipka, J. P. Watson, S. S. Oren, and R. P. O'Neill. A successive linear programming approach to solving the IV-ACOPF. *IEEE Trans. Power Systems*, 31(4):2752–2763, 2016.
- [6] J. Chen, B. Yang, and X. Guan. Optimal demand response scheduling with Stackelberg game approach under load uncertainty for smart grid. In *IEEE Int. Conference on Smart Grid Communications*, pages 546–551, 2012.
- [7] C. Coffrin and P. Van Hentenryck. A linear-programming approximation of AC power flows. *INFORMS J. Computing*, 26(4):718–734, 2014.
- [8] B. Colson, P. Marcotte, and G. Savard. Bilevel programming: A survey. *4OR: A Quarterly Journal of Operations Research*, 3:87–107, 2005.
- [9] O. Corradi, H. Ochsensfeld, H. Madsen, and P. Pinson. Controlling electricity consumption by forecasting its response to varying prices. *IEEE Trans. Power Systems*, 28(1):421–429, 2013.
- [10] B. Csáji, A. Kovács, and J. Váncza. Prediction and robust control of energy flow in renewable energy systems. In *Proceedings of the 19th IFAC World Congress*, pages 3663–3669, 2014.
- [11] S. Dempe. *Foundations of Bilevel Programming*. Kluwer, 2002.
- [12] S. Dempe and A. B. Zemkoho. The bilevel programming problem: reformulations, constraint qualifications and optimality conditions. *Mathematical Programming*, 138(1):447–473, 2013.
- [13] P. Egri, T. Kis, A. Kovács, and J. Váncza. An inverse economic lot-sizing approach to eliciting supplier cost parameters. *International Journal of Production Economics*, 149:80–88, 2014.
- [14] B. P. Esther and K. S. Kumar. A survey on residential demand side management architecture, approaches, optimization models and methods. *Renewable and Sustainable Energy Reviews*, 59:342–351, 2016.
- [15] H. Golmohamadi and R. Keypour. Stochastic optimization for retailers with distributed wind generation considering demand response. *Journal of Modern Power Systems and Clean Energy*, 6(4):733–748, 2018.
- [16] A. Gupta, R. P. Saini, and M. P. Sharma. Modelling of hybrid energy system—part I: Problem formulation and model development. *Renewable Energy*, 36(2):459–465, 2011.
- [17] M. Jalali, K. Zare, and H. Seyedi. Strategic decision-making of distribution network operator with multi-microgrids considering demand response program. *Energy*, 141:1059–1071, 2017.
- [18] A. Kovács. Bilevel programming approach to optimizing a time-variant electricity tariff for demand response. In *IEEE Int. Conference on Smart Grid Communications*, pages 674–679, 2016.
- [19] A. Kovács. On the computational complexity of tariff optimization for demand response management. *IEEE Trans. Power Systems*, 33(3):3204–3206, 2018.
- [20] A. Kovács, R. Bártai, B. Csáji, P. Dudás, B. Háty, G. Pedone, T. Révész, and J. Váncza. Intelligent control for energy-positive street lighting. *Energy*, 114:40–51, 2016.
- [21] M. Labbé and A. Violin. Bilevel programming and price setting problems. *4OR*, 11(1):1–30, 2013.
- [22] R. Lu, S. H. Hong, and X. Zhang. A dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach. *Applied Energy*, 220:220–230, 2018.
- [23] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Basar. Demand response management in the smart grid in a large population regime. *IEEE Trans. Smart Grid*, 7(1):189–199, 2016.
- [24] D. T. Nguyen, H. T. Nguyen, and L. B. Le. Dynamic pricing design for demand response integration in power distribution networks. *IEEE Trans. Power Systems*, 31(5):3457–3472, 2016.

- [25] L. Ni, W. Liu, F. Wen, Y. Xue, Z. Dong, Y. Zheng, and R. Zhang. Optimal operation of electricity, natural gas and heat systems considering integrated demand responses and diversified storage devices. *Journal of Modern Power Systems and Clean Energy*, 6(3):423–437, 2018.
- [26] Ö. Özdemir, F. D. Munoz, J. L. Ho, and B. F. Hobbs. Economic analysis of transmission expansion planning with price-responsive demand and quadratic losses by successive LP. *IEEE Trans. Power Systems*, 31(2):1096–1107, 2016.
- [27] F. Palacios-Gomez, L. Lasdon, and M. Engquist. Nonlinear optimization by successive linear programming. *Management Science*, 28(10):1106–1120, 1982.
- [28] M. Razmara, G. R. Bharati, M. Shahbakhti, S. Paudyal, and R. D. Robinett III. Bilevel optimization framework for smart building-to-grid systems. *IEEE Trans. Smart Grid*, 9(2):582–593, 2018.
- [29] P. Scott, S. Thiébaux, M. van den Briel, and P. Van Hentenryck. Residential demand response under uncertainty. In *Principles and Practice of Constraint Programming*, volume Springer LNCS 8124, pages 645–660. 2013.
- [30] I. Sharma, J. Dong, A. A. Malikopoulos, M. Street, J. Ostrowski, T. Kuruganti, and R. Jackson. A modeling framework for optimal energy management in a residential building. *Journal of Energy and Buildings*, 130:55–63, 2016.
- [31] H.M. Soliman and A. Leon-Garcia. Game-theoretic demand-side management with storage devices for the future smart grid. *IEEE Trans. Smart Grid*, 5(3):1475–1485, 2014.
- [32] W. Tushar, W. Saad, H. V. Poor, and D. B. Smith. Economics of electric vehicle charging: A game theoretic approach. *IEEE Trans. Smart Grid*, 3(4):1767–1778, 2012.
- [33] W. Tushar, J.A. Zhang, D.B. Smith, H.V. Poor, and S. Thiebaux. Prioritizing consumers in smart grid: A game theoretic approach. *IEEE Trans. Smart Grid*, 5(3):1429–1438, 2014.
- [34] M. Vallés, A. Bello, J. Reneses, and P. Frías. Probabilistic characterization of electricity consumer responsiveness to economic incentives. *Applied Energy*, 216:296–310, 2018.
- [35] 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*, 6(3):1364–1374, 2015.
- [36] M. Yu and S. H. Hong. Supply-demand balancing for power management in smart grid: A Stackelberg game approach. *Applied Energy*, 164:702–710, 2016.
- [37] M. Yu and S. H. Hong. Incentive-based demand response considering hierarchical electricity market: A Stackelberg game approach. *Applied Energy*, 203:267–279, 2017.
- [38] X. Zhang, T. Bao, T. Yu, B. Yang, and C. Han. Deep transfer Q-learning with virtual leader-follower for supply-demand Stackelberg game of smart grid. *Energy*, 133:348–365, 2017.
- [39] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen. A bilevel model for electricity retailers’ participation in a demand response market environment. *Energy Economics*, 36:182–197, 2013.