Article

An Integrative DR Study for Optimal Home Energy Management Based on Approximate Dynamic Programming

Hepeng Li [1],[†] ⬥, Peng Zeng [1],*,[†], Chuanzhi Zang [1],[†], Haibin Yu [1],[†] and Shuhui Li [2],[†]

[1] Networked Control Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; lihepeng@sia.cn (H.L.); zangcz@sia.cn (C.Z.); yhb@sia.cn (H.Y.)
[2] Department of Electrical & Computer Engineering, the University of Alabama, Tuscaloosa, AL 35487, USA; sli@eng.ua.edu
* Correspondence: zp@sia.cn; Tel.: +86-024-2397-0395
† The author contributed equally to this work.

Abstract: This paper presents an integrative demand response (DR) mechanism for energy management of appliances, an energy storage system and an electric vehicle (EV) within a home. The paper considers vehicle-to-home (V2H) and vehicle-to-grid (V2G) functions for energy management of EVs and the degradation cost of the EV battery caused by the V2H/V2G operation in developing the proposed DR mechanism. An efficient optimization algorithm is developed based on approximate dynamic programming, which overcomes the challenges of solving high dimensional optimization problems for the integrative home energy system. To investigate how the participation of different home appliances affects the DR efficiency, several DR scenarios are designed. Then, a detailed simulation study is conducted to investigate and compare home energy management efficiency under different scenarios.

Keywords: demand response; home energy management system; approximate dynamic programming; vehicle-to-home; vehicle-to-grid

1. Introduction

As a key feature of the smart grid, demand response (DR) brings reliability and efficiency to the electric system through reducing or shifting peak demand for energy [1]. Smart homes that can monitor and control their usage of electricity in real time are considered to have the greatest potential for DR [2]. Moreover, with more batteries and electric vehicles used in household environments, the potential will be greater. To fulfill the full potential, efficient DR mechanisms and advanced home energy management systems (HEMS) are crucial components and need to be studied and designed carefully [3,4].

Currently, a large body of DR research on HEMS exists, and optimal operation of home appliances considering users' electricity cost or comfort in response to a dynamic electricity price is a major concern. For example, in [5], a learning-based optimal DR policy for a heating ventilation and air conditioning system (HVAC) is developed to minimize the electricity cost. In [6], an end-users' comfort-oriented DR strategy for residential HVAC aggregation scheduling is investigated. In [7], the DR capability of electrical water heaters (EWH) is evaluated for load-shifting and balancing reserve. In [8], coordination control of multiple batteries for optimal HEMS is studied. However, these DR strategies focus on scheduling of only a single type of appliance.

Normally, different types of DR appliances and ESS could be properly scheduled to coordinate one another and improve user's benefits. Moreover, a V2H/V2G-enabled EV with bi-directional power

flow could also create synergies with smart appliances and ESS. Therefore, it is necessary to study the optimal DR strategy from an integrative perspective considering different operating features of all kinds of home appliances, ESS and V2H/V2G-enabled EV.

Although some pioneering research works about integrative DR have been performed, few provide a fully-integrated solution for an optimal DR covering all kinds of home appliances, ESS and EV. For example, in [9], coordinated control of several types of home appliances is implemented in an optimization-based DR controller, but ESS and EV are not considered. In [10], the joint operation management of different home appliances, micro-CHP and an energy storage device is optimized, but EV is not considered. In [11], various home appliances are classified into five different categories according to their operation characteristics, and an integration of the five types of home appliances for DR scheduling is investigated; however, the model of ESS and EV is highly simplified by neglecting their energy level constraints. In [12], a dynamic energy management framework considering all types of home appliances and EV is proposed, but ESS and the V2H/V2G functions of the EV are not considered. In [13], an optimal DR strategy considering the collaborative operation of an ESS and an EV with V2H/V2G functions is presented, but the DR potential of smart appliances is not exploited. In [14], different types of home appliances and an EV with V2H capability are optimally scheduled, but the coordination control between ESS and the EV battery is not presented. In the latest research [15], interestingly, the collaboration of all types of appliances, ESS and V2H/V2G-enabled EV is evaluated using a mixed-integer linear programming (MILP) DR model, which provides a meaningful reference for study in this aspect.

Nevertheless, the aforementioned studies on the integrative DR strategy only consider the consumer's electricity cost or thermal comfort and neglect the degradation cost of the EV battery. More frequent charging and discharging cycles caused by the V2H/V2G operation accelerate battery degradation and increase the wear cost [16], which has remained as a main barrier to the integration of V2H/V2G-enabled EV with household DR [17]. Therefore, whether the V2H/V2G services will help end-users save enough electricity cost to offset the additional degradation of the EV battery would have to be carefully evaluated in an integrative DR scheme.

On the other hand, a solution problem arises in developing an optimal integrative DR strategy. As more smart appliances participant in household DR and more detailed scheduling in a short time-slot framework becomes entailed, an exponentially larger optimization model is needed. The significant growth in dimension combined with complex objectives and various constraints of the appliances makes the integrative DR optimization problem very difficult to solve in limited time. However, most existing DR research works focus on the modeling process in the framework of MILP [15,18], MINLP[10], a game [19], etc., and ignores the solution problem by tailoring the model to suit a certain commercial solver. Other research on the DR algorithm [4] has also not been found to involve the discussion about how to handle the rapid growth in dimensions with the increase of the number of controllable appliances.

Approximate dynamic programming (ADP) [20] is a powerful tool for high dimensional function optimization problems, which has been examined by some previous works [8,21–24]. For example, in [8,23,24], ADP is used to solve the optimal battery management and multi-battery coordination control problem in smart home environments. However, the integrated optimization of all kinds of smart home appliances, the energy storage system and EV, which includes high dimensional continuous and integer variables in the framework of ADP, has not been found in the literature.

In this paper, an integrative DR study on optimal scheduling of different types of appliances, ESS and a V2H/V2G-enabled EV considering the battery degradation cost of the EV is presented. A solution method based on ADP is developed for the integrative DR optimization problem. In the developed method, a polynomial function for optimal value function approximation is designed to suit the problem. The contributions of the paper include: (1) the formulation of a joint optimization of all types of residential appliances, ESS and EV considering the electricity consumption cost, user's thermal comfort and battery degradation of the EV; (2) a solution method based on the design of ADP for the

integrative DR scheduling to overcome the difficulty of solving the high dimensional optimization problem due to the increasing number of DR-capable appliances; (3) detailed comparative analysis about how the integration of different home appliances affects the DR efficiency.

The rest of the paper is organized as follows. Section 2 gives the framework of the HEMS and models of appliances, energy storage system and EV for DR optimization in different scenarios. In Section 3, the proposed optimization algorithm is presented. The case studies are given in Section 4. Section 5 draws the conclusions.

2. Modeling of the Integrative DR Strategy

In this section, the integrative DR problem is formulated aiming for optimal day-ahead scheduling of all DR appliances within a home. A suitable cyber infrastructure is presumptively equipped to collect appliances' parameters, end-user's preferences and to receive day-ahead hourly electricity price signals from the utility. Weather information about ambient temperature and solar radiation is accessed through the Internet or forecasted through a local forecasting system in the HEMS. Other parameters such as arrival/departure time of the EV, hot water usage and critical appliances' consumption are considered to be known by occupants' habit statistics.

2.1. Models and Constraints of Home Appliances

In an existing home, primary home appliances include the heating/cooling, ventilation and air conditioning system (HVAC), electric water heating (EWH), clothes washing (CW) machine, clothes dryer (CD), dishwasher (DW), cooking, lighting, entertainments, etc. Typically, these appliances are classified into three categories: critical appliances, shiftable appliances and adjustable appliances (Figure 1). Critical appliances are DR incapable. Shiftable appliances can be shifted, but cannot be regulated. Adjustable appliances can be fully controlled and adjusted. In general, the HVAC and the EWH are adjustable appliances. The CW, CD and DW are shiftable appliances, and the rest are considered as critical. For a future home, additional main home appliances would include an energy storage system (ESS), photovoltaic (PV) panels and electric vehicles. Next, the operating constraints associated with these home appliances are formulated for the integrative DR problem.
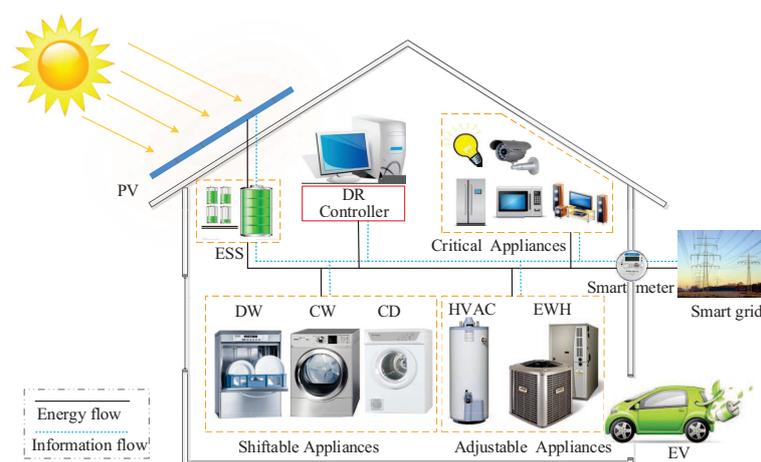
Figure 1. Categories of home appliances. DR, demand response; DW, dishwasher; CW, clothes washer; CD, clothes drier; EWH, electric water heater.

2.1.1. Adjustable Appliances

Adjustable appliances include HVAC and EWH, whose power consumption is flexibly controlled to provide thermal comfort to the occupants.

For the HVAC, the indoor temperature is maintained at a comfortable level according to the occupants' preference setting. Temperature deviation is allowed if the occupants are willing to sacrifice their comfort for a lower electricity bill. A thermal model from [25] is adopted to describe the dynamics of the indoor temperature as given below:

$$T_{i+1,in}^{HVAC} = \delta \cdot T_{i,in}^{HVAC} + (1-\delta)(T_{i,out}^{HVAC} - \eta^{HVAC} \cdot \frac{P_i^{HVAC} \cdot \Delta t}{A}) \tag{1}$$

$$T_{set}^{HVAC} - \Delta T^{HVAC} \leq T_{i,in}^{HVAC} \leq T_{set}^{HVAC} + \Delta T^{HVAC} \tag{2}$$

$$0 \leq P_i^{HVAC} \leq P_{max}^{HVAC} \tag{3}$$

where in (1), $T_{i,in}^{HVAC}$ is the indoor temperature in time slot i, $T_{i,out}^{HVAC}$ is the outdoor temperature in time slot i, δ is the inertia factor, η is the efficiency of the HVAC and A is the thermal conductivity (kW/°F) of the house. In (2), T_{set}^{HVAC} denotes the temperature setting, and ΔT^{HVAC} is the allowable deviation. The average power (kW) drawn by the HVAC in time slot i is denoted as P_i^{HVAC}, which is constrained by its upper limit P_{max}^{HVAC} (kW).

Similarly, the EWH regulates its power consumption to keep the water temperature within a comfortable range. Considering the inlet cold water replenished into the EWH and the heat exchange with the indoor air, a thermal model from [26] is used to describe the dynamics of the hot water temperature as shown in (4). The operational constraints of the EWH are given below.

$$T_{i+1,in}^{EWH} = e^{-\left(\frac{\Delta t}{R' \cdot K}\right)} \cdot T_{i,in}^{EWH} + \left(1 - e^{-\left(\frac{\Delta t}{R' \cdot K}\right)}\right) \cdot [G \cdot T_{i,in}^{HVAC} + B_i \cdot T_{cold}^{EWH} + Q_i^{EWH}] \cdot R' \tag{4}$$

$$G = \frac{SA}{R}, \quad B_i = d_{water} \times F_i \times C_p, \quad R' = \frac{1}{G+B_i} \tag{5}$$

$$K = volumn \times d_{water} \times C_p, \quad Q_i^{EWH} = 3412.1 \times P_i^{EWH} \times \Delta t \tag{6}$$

$$T_{set}^{EWH} - \Delta T^{EWH} \leq T_{i,in}^{EWH} \leq T_{set}^{EWH} + \Delta T^{EWH} \tag{7}$$

$$0 \leq P_i^{EWH} \leq P_{max}^{EWH} \tag{8}$$

where in (4)–(6), $T_{i,in}^{EWH}$ denotes the hot water temperature (°F) in time slot i, T_{cold}^{EWH} is the inlet cold water temperature (°F), SA is the tank surface area (ft^2), R is the tank insulation thermal resistance (hour·ft^2·°F/BTU), d_{water} is the density of water (8.34 lbs/gallon), F_i is the hot water flow rate (gallons/hour) in time slot i, C_p is the specific heat of water (1.00 BTU/(lbs·°F)), *volumn* is the capacity of the tank (gallons) and G, B_i, R', K and Q_i^{EWH} are intermediate variables. In (7), T_{set}^{EWH} is the temperature setting, and ΔT^{EWH} is the allowable tolerance. In (8), the average power P_i^{EWH} in time slot i is constrained by its upper limit P_{max}^{EWH}.

2.1.2. Shiftable Appliances

Considered shiftable appliances include CW, CD and DW. Taking the CW as an example, the constraints of shiftable appliances are modeled below.

Assume that a CW requires continuous operation of J^{CW} time slots to fulfill its task, and it must be finished in a time interval $[i_\alpha^{CW}, i_\alpha^{CW} + K^{CW} - 1]$ with $K^{CW}(K^{CW} > J^{CW})$ time slots. Let s_i^{CW} indicate the status of the CW in time slot i, which equals one if the CW is ON, whereas zero if the CW is OFF. Then, we have,

$$s_i^{CW} = 0, \ i \notin [i_\alpha^{CW}, i_\alpha^{CW} + K^{CW} - 1] \tag{9}$$

$$s_i^{CW} = 1, \text{if } s_{i-1}^{CW} = 1 \text{ and } \sum_i s_{i-1}^{CW} < J^{CW} \tag{10}$$

$$\sum_i s_i^{CW} = J^{CW}, \ i \in [i_\alpha^{CW}, i_\alpha^{CW} + K^{CW} - 1] \tag{11}$$

where (9) constrains the operation time of the CW, (10) keeps the CW running continuously and (11) ensures that the task of the CW is fulfilled.

The power consumption P_i^{CW} of the CW in time slot i is determined by its status s_i^{CW} and power patterns. Let $P_{(j)}^{CW}$ be the power pattern of the CW at a specific operating sequence $j \in \{1, 2, \ldots, J^{CW}\}$, then P_i^{CW} can be obtained as below.

$$P_i^{CW} = \begin{cases} 0, & s_i^{CW} = 0; \\ P_{(j)}^{CW}, & \sum_i s_i^{CW} = j \text{ and } s_i^{CW} \neq 0; \end{cases} \tag{12}$$

The models of CD and DW can be obtained in a similar way.

2.1.3. Energy Storage

Let the power of the ESS be positive if the ESS is charging and negative if the ESS is discharging. Therefore, the model of the ESS can be described using the state of charge (SOC) as below,

$$SOC_i^{ESS} = \begin{cases} SOC_{i-1}^{ESS} + \frac{P_i^{ESS} \cdot \Delta t}{E_{max}^{ESS}} \cdot \eta_{ch}^{ESS}, & P_i^{ESS} \geq 0; \\ SOC_{i-1}^{ESS} + \frac{P_i^{ESS} \cdot \Delta t}{E_{max}^{ESS}} \cdot \frac{1}{\eta_{dis}^{ESS}}, & P_i^{ESS} < 0. \end{cases} \tag{13}$$

$$-P_{dis,max}^{ESS} \leq P_i^{ESS} \leq P_{ch,max}^{ESS} \tag{14}$$

$$SOC_{min}^{ESS} \leq SOC_i^{ESS} \leq SOC_{max}^{ESS} \tag{15}$$

where SOC_i^{ESS} is the SOC of the ESS at the end of the time slot i and P_i^{ESS} is the power of the ESS (kW) in time slot i, which is limited by the maximum discharging power $P_{dis,max}^{ESS}$ and maximum charging power $P_{ch,max}^{ESS}$ of the ESS; E_{max}^{ESS} is the maximum capacity of the ESS (kWh), and η_{ch}^{ESS} and η_{dis}^{ESS} are the charging and discharging efficiency of the ESS, respectively; the SOC of the ESS is constrained between the allowable minimum SOC_{min}^{ESS} and the allowable maximum SOC_{max}^{ESS}.

2.1.4. Electric Vehicle

In modeling of an EV, the enhanced V2G and V2H functions of the EV are considered. These functions are active only when the EV is at home. Similar to the ESS, let the EV power be positive if it is charging and negative if it is discharging. The model of the EV is described in (16)–(18).

$$SOC_i^{EV} = \begin{cases} SOC_{i-1}^{EV} + \frac{P_i^{EV} \cdot \Delta t}{E_{max}^{EV}} \cdot \eta_{ch}^{EV}, & P_i^{EV} \geq 0; \\ SOC_{i-1}^{EV} + \frac{P_i^{EV} \cdot \Delta t}{E_{max}^{EV}} \cdot \frac{1}{\eta_{dis}^{EV}}, & P_i^{EV} < 0. \end{cases} \tag{16}$$

$$-P_{d\_max}^{EV} \leq P_i^{EV} \leq P_{ch\_max}^{EV} \tag{17}$$

$$SOC_{min}^{EV} \leq SOC_i^{EV} \leq SOC_{max}^{EV} \tag{18}$$

The meanings of the terms in (16)–(18) are similar to those of the ESS model (Equations (13)–(15)).

Moreover, the EV has to be fully charged before it leaves home. Assume the EV arrives home at the beginning of the time slot i_α^{EV} and is available at home during the time interval $[i_\alpha^{EV}, i_\beta^{EV}]$, so the SOC of the EV should meet the constraints,

$$SOC_{i_\alpha^{EV}-1}^{EV} = SOC_{max}^{EV} - \frac{E_{driven}^{EV} / \eta_{dis}^{EV}}{E_{max}^{EV}} \tag{19}$$

$$SOC_{i_\beta^{EV}}^{EV} = SOC_{max}^{EV} \tag{20}$$

where $E_{driven}^{EV} = d_d / \eta_{driven}^{EV}$ is the total energy consumption for driving the car, d_d is the driven distance that is assumed to be predictable and η_{driven}^{EV} is the driving efficiency representing energy needed to drive an EV per mile.

2.2. DR Optimization Scenarios

To investigate how the integration of different home appliances affects the DR efficiency, four DR optimization scenarios are designed.

2.2.1. Optimal DR for Adjustable Appliances

The first scenario assumes that only the adjustable appliances, i.e., the HVAC and EWH, participant in the DR scheme and that the other appliances act as critical appliances. The ESS and PV system are not available. Then, the DR model for this scenario is formulated as below,

$$
\begin{aligned}
(P1) \ \min \ J_1 &= \sum_{i=1}^{I} price_i \cdot P_i^{total} \cdot \Delta t \\
\text{s.t.} \ \ P_i^{total} &= P_i^{adj} + P_i^{shi} + P_i^{cri} + P_i^{EV} \\
P_i^{adj} &= P_i^{HVAC} + P_i^{EWH} \\
P_i^{shi} &= P_i^{CW} + P_i^{CD} + P_i^{DW} \\
&\text{and (1)–(8)}
\end{aligned}
\tag{21}
$$

where $price_i$ is the electricity price in time slot i and P_i^{total} is the total power consumption, which equals the sum of the power consumption of adjustable appliances P_i^{adj}, shiftable appliances P_i^{shi}, critical appliances P_i^{cri} and EV P_i^{EV}.

Based on the assumption in this scenario, the power consumptions of shiftable appliances, critical appliances and EV are all constant, so removing them from the objective function makes no difference for the optimal solution. Therefore, the problem (P1) can be simplified as (22).

$$
\begin{aligned}
\min \ &\sum_{i=1}^{I} price_i \cdot (P_i^{HVAC} + P_i^{EWH}) \cdot \Delta t \\
\text{s.t.} \ \ &(1)–(8)
\end{aligned}
\tag{22}
$$

2.2.2. Optimal DR Policy Considering Shiftable Appliances

Second, we consider the scenario that shiftable appliances also participate in DR along with adjustable appliances. In this scenario, additional discrete variables need to be solved to determine the status of the CD, CW and DW; therefore, a mixed integer programming problem (P2) is formed and can be expressed in an equivalent format as shown in (23).

$$
\begin{aligned}
(P2) \ \min \ J_2 &= \sum_{i=1}^{I} price_i \cdot P_i^{total} \cdot \Delta t \\
\text{s.t.} \ \ P_i^{total} &= P_i^{adj} + P_i^{shi} + P_i^{cri} + P_i^{EV} \\
\iff \ \min \ &\sum_{i=1}^{I} price_i \cdot (P_i^{adj} + P_i^{shi}) \cdot \Delta t \\
\text{s.t.} \ (1)&–(12)
\end{aligned}
\tag{23}
$$

2.2.3. Optimal DR Policy Combining ESS and PV

Third, we consider the scenario that the energy storage system and PV panels are available. The DR optimization problem for this scenario can be modeled as (P3).

$$(P3) \quad \min \ J_3 = \sum_{i=1}^{I} price_i \cdot P_i^{total} \cdot \Delta t$$

$$\text{s.t.} \ P_i^{total} + P_i^{pv} = P_i^{applis} + P_i^{ESS} + P_i^{EV} \tag{24}$$

$$P_i^{applis} = P_i^{adj} + P_i^{shi} + P_i^{cri}$$

$$\text{and (1)–(15)}$$

where P_i^{pv} is the power production of PV in time slot i.

2.2.4. Integrative Optimal DR Policy

In the last scenario, the V2H/V2G applications of the EV are enabled. Since the V2H/V2G operation may accelerate battery aging and shorten the cycle life, the degradation cost of the EV battery is considered in the formulation. To develop an integrative optimal DR policy for this scenario, a mixed integer nonlinear programming model is needed,

$$(P4) \quad \min \ J_4 = \sum_{i=1}^{I} (price_i \cdot P_i^{total} + C_{i,deg}^{EV} \cdot |P_{i,dis}^{EV}|) \cdot \Delta t$$

$$\text{s.t.} \ P_i^{total} + P_i^{pv} = P_i^{applis} + P_i^{ESS} + P_i^{EV} \tag{25}$$

$$P_i^{applis} = P_i^{adj} + P_i^{shi} + P_i^{cri}$$

$$\text{and (1)–(20)}$$

where $P_{i,dis}^{EV}$ denotes the power discharged from the EV battery to the home appliances or to the grid, $C_{i,deg}^{EV}$ represents the degradation cost of the EV in the time slot i due to the V2G/V2H operation, which can be modeled as a function of the actual battery cycle life as below,

$$C_{i,deg}^{EV} = \frac{C_{capital}^{EV}}{L_{i,E}^{EV}} \tag{26}$$

where $C_{i,capital}^{EV}$ is the capital cost of the EV battery ($/kWh) and $L_{i,E}^{EV}$ is the battery life of the EV throughput energy (kWh). The battery life (kWh) can be expressed as below,

$$L_{i,E}^{EV} = E_{max}^{EV} \cdot L_{i,N}^{EV} = E_{max}^{EV} \cdot f(DoD_i^{EV}) \tag{27}$$

where $L_{i,N}^{EV}$ represents the battery life in number of cycles, which is a function of the depth of discharging (DoD) depending on the type of the battery. In our study, a linear function relationship between cycle life and DoD is used as below [27],

$$f(DoD_i^{EV}) = a \cdot DoD_i^{EV} + b \tag{28}$$

where $a = -4775$ and $b = 4995$. The DoD of the EV battery in time slot i can be estimated as below [27],

$$DoD_i^{EV} = \frac{E_{driven}^{EV} + (\sum_i (|P_{i,dis}^{EV}| \cdot \Delta t)/\eta_{dis}^{EV})}{E_{max}^{EV}} \tag{29}$$

3. Approximate Dynamic Programming

Solving the integrative DR optimization problem faces some challenges. First, it is a complex MINLP problem. Second, the dimension of decision variables grows rapidly as more DR-capable appliances are integrated, such as from (P1)–(P4). Third, the growth in dimension is multiplied as the time granularity gets reduced. In this section, the ADP technique is introduced to the integrative DR optimization problem to overcome the challenges. In particular, a polynomial function architecture is designed to approximate the optimal value function. Then, using the approximate optimistic policy iteration (AOPI), an optimal DR policy is derived.

3.1. Problem Reformulation

We reformulate the optimization problems in the DP term and show the reformulation process with respect to the problem (P4) for the explanation. Let $S_i = (T_{i,in}^{HVAC}, T_{i,in}^{EWH}, SOC_i^{ESS}, SOC_i^{EV}, \sum_i s_i^{CW}, \sum_i s_i^{CD}, \sum_i s_i^{DW})^{\mathrm{T}}$ be the system state at the end of time slot i, which is a vector including indoor temperature $T_{i,in}^{HVAC}$, hot water temperature $T_{i,in}^{EWH}$, the SOC of ESS SOC_i^{ESS}, the SOC of EV SOC_i^{EV} and how many time slots the CW, the CD and the DW have been powered, denoted by $\sum_i s_i^{CW}, \sum_i s_i^{CD}$ and $\sum_i s_i^{DW}$, respectively. Let $x_i = (P_i^{HVAC}, P_i^{EWH}, P_i^{ESS}, P_i^{EV}, s_i^{CW}, s_i^{CD}, s_i^{DW})^{\mathrm{T}}$ be the decision vector in time slot i, where $P_i^{HVAC}, P_i^{EWH}, P_i^{ESS}, P_i^{EV}$ are continuous variables and $s_i^{CW}, s_i^{CD}, s_i^{DW}$ are discrete ones. S_i and x_i take values in their feasible sets $\mathcal{S}$ and $\mathcal{X}_i$, respectively, which are defined by the constraints in (P4). Let $a_i : S_i \to x_{i+1}, a_i \in \mathcal{A}_i$ be a mapping from the current state to a decision, where $\mathcal{A}_i$ is the set of all feasible mappings while in S_i. Let $C_{i+1}(S_i, a_i)$ be the cost applying a_i while in S_i and $V_{i+1}(S_{i+1})$ be the total minimum cost for the residual time slots in S_{i+1} (Figure 2) and $V_I(S_I) \equiv 0$. For convenience, we redefine $i \in \{0, 1, \ldots, I-1\}$, and the initial system state is denoted by S_0. The system evolves via transition function $S_{i+1} = \mathbf{S}(S_i, a_i(S_i))$, which can be obtained from the models in Section 2. Based on Bellman's principle of optimality [28], the optimal decisions of (P4) can be obtained by solving the following Bellman equations recursively.

$$V_i(S_i) = \min_{a_i \in \mathcal{A}_i} \{C_{i+1}(S_i, a_i) + V_{i+1}(S_{i+1})\}, \forall i$$

$$J_4 = V_0(S_0)$$

(30)

In this way, we avoid solving a large optimization problem. However, the classical DP algorithm, which requires finding $V_i(S_i)$ at every S_i, and thus, suffers from the curse of dimensionality when a system includes large or continuous state and action spaces, cannot be used in our problems.
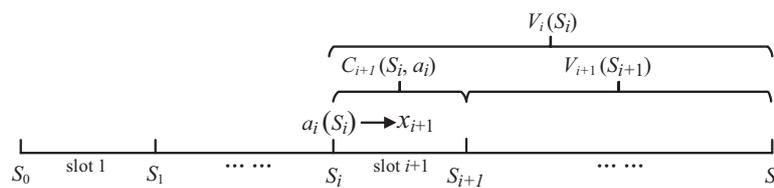

Figure 2. Relationship among notations and time slots.

3.2. ADP for the Integrative DR

To overcome the challenge, ADP is introduced to the integrative DR optimization problem. ADP breaks the curse of dimensionality by using an approximation $\widetilde{V}_i$ to the value function (or cost-to-go function) V_i, which only requires values of $V_i(S_i)$ at some states for regression. It proceeds in an iterative way and forward through time. For the formulated problems, an ADP algorithm based on approximate optimistic policy iteration (AOPI) is designed.

Assume that we start with a policy $\pi^{(n)} = \{a_0^{(n)}, a_1^{(n)}, \ldots, a_{I-1}^{(n)}\}, \forall a_i^{(n)} \in \mathcal{A}_i$ ($n \in \mathbf{N}$, here $n = 0$), and we evaluate the corresponding value function $V_i^{(n)}$ using a parametric approximator $\widetilde{V}_i(S_i | \boldsymbol{\theta}_i)$ for

all i, where $\widetilde{V}$ is the approximation architecture and $\boldsymbol{\theta}$ is a parameter vector. This process is called policy evaluation. Once $\boldsymbol{\theta}$ is determined, we can obtain an approximating value function $\widetilde{V}_i(S_i|\boldsymbol{\theta}_i^{(n)})$ for all i with respect to policy $\pi^{(n)}$. Then, we use $\widetilde{V}_i(S_i|\boldsymbol{\theta}_i^{(n)})$ to obtain an improved policy $\pi^{(n+1)} = \{a_0^{(n+1)}, a_1^{(n+1)}, \ldots, a_{I-1}^{(n+1)}\}$ by computing the Bellman functions as below.

$$a_i^{(n+1)} = \arg\min_{a_i \in \mathcal{A}_i} \{C_{i+1}(S_i, a_i) + \widetilde{V}_{i+1}(S_{i+1}|\boldsymbol{\theta}_{i+1}^{(n)})\} \tag{31}$$

After that, the corresponding value function $\widetilde{V}_i(S_i|\boldsymbol{\theta}_i^{(n+1)})$ is evaluated again, and we repeat the procedure until a converged parameter sequence $\{\boldsymbol{\theta}_i^{(0)}, \boldsymbol{\theta}_i^{(1)}, \ldots\} \mapsto \boldsymbol{\theta}_i^{(*)}$ is found for all i.

Next, we show the process of policy evaluation. First, in each iteration n, we initialize the parameter vector with $\boldsymbol{\theta}_i^{(n,0)}$ (for $n \geq 1, \boldsymbol{\theta}_i^{(n,0)} = \boldsymbol{\theta}_i^{(n-1)}$). Then, we choose a start state S_0 and calculate one period cost $C_{i+1}(S_i, a_i^{(n)})$ for all i forward in time based on the transform function $S_{i+1} = \mathbf{S}(S_i, a_i^{(n)}(S_i))$. After that, we calculate the cumulative cost to obtain an observation of the value function $v_i = \sum_i^{I-1} C_{i+1}(S_i, a_i^{(n)})$ for all i. With the pair of (S_i, v_i), we compute the new parameter vector $\boldsymbol{\theta}_i^{(n,1)}$ for all i via some learning algorithm. Repeat the procedure a certain number of times, say M, and then, we assign $\boldsymbol{\theta}_i^{(n,M)}$ to $\boldsymbol{\theta}_i^{(n)}$ for approximating value function $\widetilde{V}_i(S_i|\boldsymbol{\theta}_i^{(n)})$ for all i.

3.3. Approximating Functions' Design

Obviously, the approximation architecture and learning algorithm are significant for the convergence and performance of the algorithm. In this study, a linear parameter architecture-based multivariate polynomial is designed for the value function approximation as shown in (32),

$$\widetilde{V}_i(S_i|\boldsymbol{\theta}_i) = \theta_{i,0} + \sum_{k=1}^{7} \theta_{i,k} S_i(k) + \sum_{r=2}^{Q}(\sum_{k=1}^{7} \theta_{i,k^r} S_i(k)^r + \sum_{q=1}^{r-1} \sum_{k=1}^{6} \sum_{l=k+1}^{7} \theta_{i,k^{r-q}l^q} S_i(k)^{r-q} S_i(l)^q) \tag{32}$$

where $\boldsymbol{\theta}_i = (\theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,F})^{\mathrm{T}}$ is the parameter vector to be estimated, $S_i(k)$ is the k-th element of the state vector S_i and $r, q \in \mathbf{N}$ are nonnegative integer powers; here, Q is the degree of the polynomial.

The polynomial function is adopted for the value function approximation because of its many merits. First, it is a linear parameter architecture that is easy to train and has fast convergence speed. Second, it can approximate closely to any nonlinear function on a finite interval according to the Weierstrass approximation theorem [29]. Third, it is differentiable and simple, which reduces the complexity of policy evaluation and the proposed method in general. In this study, via trial and error, a cubic polynomial containing 1, 2, 3 power-terms of each state variable, two-degree cross-terms and a constant term is chosen to approximate the value function as shown in (33).

$$\widetilde{V}_i(S_i|\boldsymbol{\theta}_i) = \sum_{k=1}^{7} \sum_{n=1}^{3} \theta_{i,k^r} S_i(k)^r + \sum_{k=1}^{7} \sum_{l=k+1}^{7} \theta_{i,kl} S_i(k) S_i(l) + \theta_{i,0} \tag{33}$$

Given the approximation architecture, the recursive least squares algorithm (RLS) [30], which is a quickly-converged learning algorithm for the linear regression problem, is applied to update the parameter vector in each policy evaluation step. It should be noted that the policy improvement step in (31) also plays an important role in the performance of the algorithm. In fact, it is hard to obtain the improved policy exactly because it involves solving a mixed integer nonlinear optimization problem. The pseudo-code of AOPI is shown in Algorithms 1 and 2.

Algorithm 1. AOPI.

Input:
   1. Approximate function $\widetilde{V}_i(S_i|\boldsymbol{\theta}_i)$;
   2. Initial policy $\pi^{(0)} = \{a_0^{(0)}, a_1^{(0)}, \ldots, a_{I-1}^{(0)}\}, \forall a_i^{(0)} \in \mathcal{A}_i$;
   3. $n \leftarrow 0$;
Output: Optimal parameter vector θ_i^* for all i; optimal decision policy $\pi^{(*)}$;
   4. **while** $\sum_i^I \|\theta_i^{(n)} - \theta_i^{(n-1)}\| < \varepsilon(n > 1)$ **do**
   5.   **Algorithm 2** Policy evaluation;
   6.   Obtain parameter vector $\theta_i^{(n)}$ for all i;
   7.   Solve (31);
   8. **end while**;
   9. **return** $\theta_i^{(*)}$ for all i.

Algorithm 2. Policy evaluation.

Input: $n, \pi, \theta_i^{(n-1)}$ for all i;
Output: $\theta_i^{(n)}$ for all i;
   1. Initialization: $\theta_i^{(n,0)} \leftarrow \theta_i^{(n-1)}$(if $n = 0, \theta_i^{(n,0)} \leftarrow 0$)
   2. **for** $m = 1$ to M **do**
   3.   Sample an initial state $S_0^{(m)}$;
   4.   **for** $i = 0$ to $I - 1$ **do**
   5.   Compute $C_{i+1}^{(m)}(S_i^{(m)}, a_i^{(n)})$;
   6.   Compute $S_{i+1}^{(m)} = \textbf{TS}(S_i^{(m)}, a_i^{(n)})$;
   7.   **end for**
   8.   **for** $i = 0$ to $I - 1$ **do**
   9.   Compute $v_i^{(m)} = \sum_i^{I-1} C_{i+1}(S_i^{(m)}, a_i^{(n)})$;
   10.   Update $\theta_i^{(n,m-1)}$ to $\theta_i^{(n,m)}$ by RLS;
   11.   **end for**
   12. **end for**;
   13. **return** $\theta_i^{(n)} = \theta_i^{(n,M)}$.

---

4. Case Studies

4.1. Parameters

In this section, numerical simulation is implemented to test the optimal DR policies and the proposed ADP algorithm. Table 1 summarizes all appliance parameters and characteristics. The real outdoor temperature on a hot day [31] in June 2014 in the U.S. state of Illinois and the statistical hot water flow rate [32] and aggregated critical loads [33] on the same date are adopted. The PV production data from [33] and the Ameren Illinois' DAPtariffs [34] on a high DAP day in June 2014 are used in the simulation. Figure 3 depicts the used data. We assume the DR policy starts from 8:00 a.m. and runs for 24 h with 15-min intervals. Choose the convergence tolerance $\varepsilon = 10^{-3}$. The simulations platform is MATLAB 2014a in an Intel(R) Core(TM) i5-3475S, 2.90-GHz personal computer with 8 GB RAM memory.

The developed integrative DR optimization model is a non-convex MINLP and belongs to a class of NP-hard problems, and no algorithm guarantees a global optimum solution in polynomial time. In order to examine the effectiveness of the designed ADP algorithm, we compare it with two existing non-convex MINLP technologies: one is a standard integer branch and bound algorithm developed by the YALMIP Optimization toolbox BNB solver [35], in which the associated relaxed problem is solved by the MATLAB build-in nonlinear constrained optimization solver 'fmincon'; the other is genetic algorithm (GA) provided by the Global Optimization Toolbox in MATLAB. For GA, we choose the population size to be 50; choose the elite population to be 0.05-times the population size and

the crossover fraction to be 0.8; default selection and mutation operators in MATLAB are used; the algorithm stops when the average relative change in the best fitness function value generations is less than or equal to 10^{-20}; the maximum generation is set to be 100-times the number of populations.
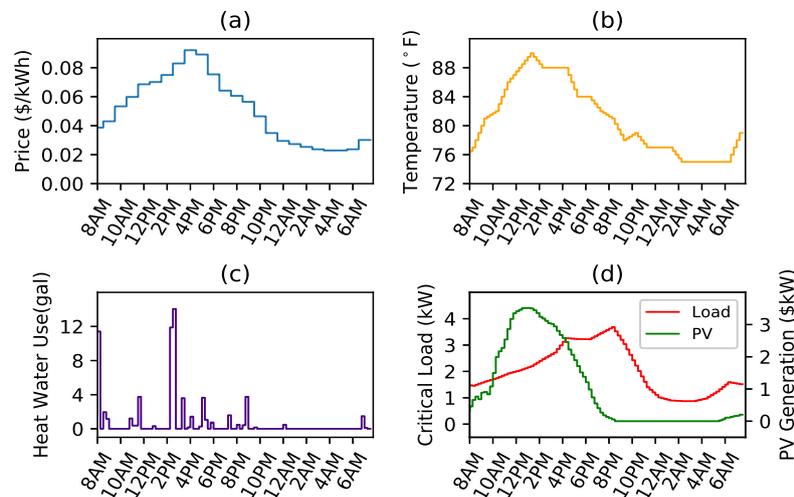

Figure 3. Data: (**a**) DAPtariffs. (**b**) Outdoor temperature. (**c**) Hot water flow. (**d**) Critical load and PV production.

Table 1. Parameters of the residential appliances.

HVAC						
T_{set}^{HVAC}	ΔT^{HVAC}	$T_{0,in}^{HVAC}$	P_{max}^{HVAC}	δ	η	A
73 °F	2 °F	73 °F	4 kW	0.95	3	0.25

EWH							
T_{set}^{EWH}	ΔT^{EWH}	$T_{0,in}^{EWH}$	T_{cold}^{EWH}	P_{max}^{EWH}	SA	A	$volumn$
125 °F	5 °F	125 °F	60 °F	4.5 kW	24.1 ft^2	15	40 gal

ESS					
E_{max}^{ESS}	SOC_0^{ESS}	SOC_{max}^{ESS}	SOC_{min}^{ESS}	P_{ch/d_max}^{ESS}	$\eta_{ch/d}^{ESS}$
5 kWh	0.6	1	0.2	1 kW/1kW	0.95/0.95

EV				
E_{max}^{EV}	SOC_{max}^{ESS}	SOC_{min}^{ESS}	$P_{ch\_max}^{ESS}$	$P_{ch\_max}^{ESS}$
21.6 kWh	1	0.15	3 kW	3 kW
$\eta_{ch/d}^{ESS}$	$[i_\alpha^{EV}, i_\beta^{EV}]$	$C_{capital}^{EV}$	η_{driven}^{EV}	d_d
0.95	[46,96]	211.9$/kWh	5.6 miles/kWh	25.68 miles

CW/DW/CD	
$\{P_{(1)}^{CW}, \dots, P_{(J^{CW})}^{CW}\}$	$[i_\alpha^{CW}, i_\alpha^{CW} + K^{CW} - 1]$
{0.5 kW, 0.5 kW, 0.5 kW, 0.5 kW}	[5, 40] → 9:00 a.m.–6:00 p.m.
$\{P_{(1)}^{DW}, \dots, P_{(J^{DW})}^{DW}\}$	$[i_\alpha^{DW}, i_\alpha^{DW} + K^{DW} - 1]$
{1 kW, 1 kW, 1 kW, 1 kW}	[7, 36] → 9:30 a.m.–5:00 p.m.
$\{P_{(1)}^{CD}, \dots, P_{(J^{CD})}^{CD}\}$	$[i_\alpha^{CD}, i_\alpha^{CD} + K^{CD} - 1]$
{4 kW, 4 kW, 4 kW, 4 kW, 4 kW, 4 kW}	[41, 96] → 6:00 p.m.–8:00 p.m.

Figure 4 shows the comparison of the daily cost and running time by using the three algorithms under different DR scenarios. The detailed comparison results are listed in Table 2.
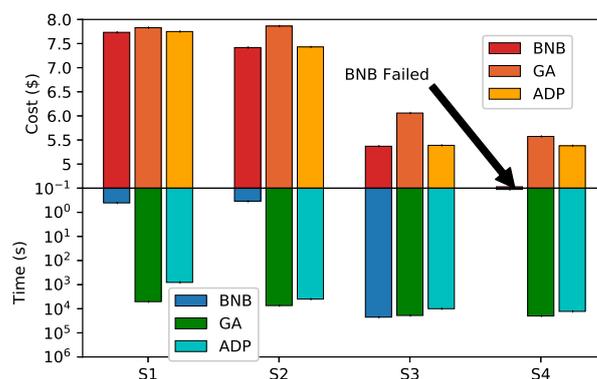
Figure 4. Comparison among different algorithms under Scenarios 1, 2, 3 and 4. ADP, approximate dynamic programming.

Table 2. Comparison of the total cost and CPU time with different solvers for the residential DR problems.

	P1		P2		P3		P4	
	cost ($)	time (s)	cost ($)	time (s)	cost ($)	time (s)	cost ($)	time (s)
BNB	7.7333	0.4	7.4176	0.4	5.3708	22734	-	-
1-9 GA *	7.8293	5109	7.8633	7431	6.0596	19,104	5.5741	20,174
ADP	7.7497	808	7.4304	4032	5.3807	10,142	4.9773	12,481

* Mean value of 10 runs.

It can be seen from Figure 4 that for the S1 and S2 DR scenarios, the BNB algorithm performs the best in terms of both the running time and total electricity cost, and the GA performs the worst; the ADP algorithm obtains a comparable result to the BNB algorithm in terms of the total cost, but consumes much more computational time. This is because the optimization models associated with the S1 and S2 DR scenarios are simple, which are a linear model for S1 and a mixed integer linear model for S2, and have a relatively small number of variables. Therefore, the BNB algorithm is able to find the global optimal minimum within a very short time. However, the ADP algorithm needs to solve the Bellman equation at each iteration and do regression to approximate the value function, so it requires more time for calculation and obtains a sub-optimal solution due to the approximation error.

Nevertheless, for more integrative DR scenarios, i.e., S3 and S4, the advantage of the ADP algorithm appears. From Figure 4, it can be seen that the ADP algorithm obtains a close or relatively low total electricity cost in a shorter time for the S3 and S4 DR scenarios compared to the BNB and GA algorithms. Notice that the scenarios S3 and S4 complicate the DR optimization model by not only enlarging its size with more continuous and integer variables, but also introducing a large number of nonlinear equality and inequality constraints, including the SOC dynamics equation of the ESS and EV charging demand constraints. This significantly burdens the BNB algorithm in the branch process due to the large number of integer variables and the bound process due to the serious nonlinearity of the relaxed problem. As a result, it fails to give feasible results after running 24 h. These complex constraints also give rise to a problem for the GA algorithm in searching feasible and better solutions wile evolving from generation to generation. Although the ADP algorithm also encounters these problems, it breaks the large nonlinear mixed integer programming model into many small optimization problems by approximating the value function, which are easy and fast to solve. Therefore, the growth in computational time of the ADP is not as fast as the BNB and the GA algorithms. Observation of the comparison suggests that the proposed ADP algorithm is effective and has special strength for complicated high dimensional optimization problems.

Figure 5 shows the optimal DR solutions of problem S4 by using the ADP algorithm. As can be seen, for the HVAC and EWH, there is a clear pre-cooling and preheating operation respectively during low price hours to reduce power consumption in high price hours, but there is still oscillating power consumption by the EWH during peak hours from 3 p.m.–10 p.m. matching the fluctuation in the usage of hot water as shown in Figure 3c in order to keep the hot water temperature at a comfortable level; the energy usages of CW, CD and DW are also shifted to the time when the price is low; the ESS is controlled to charge up from the PV or the grid in low price hours and discharge to power the loads in high price hours; the EV battery also supplies power to home appliances during peak load hours 9 p.m.–11 p.m. via the V2G/V2H operation.
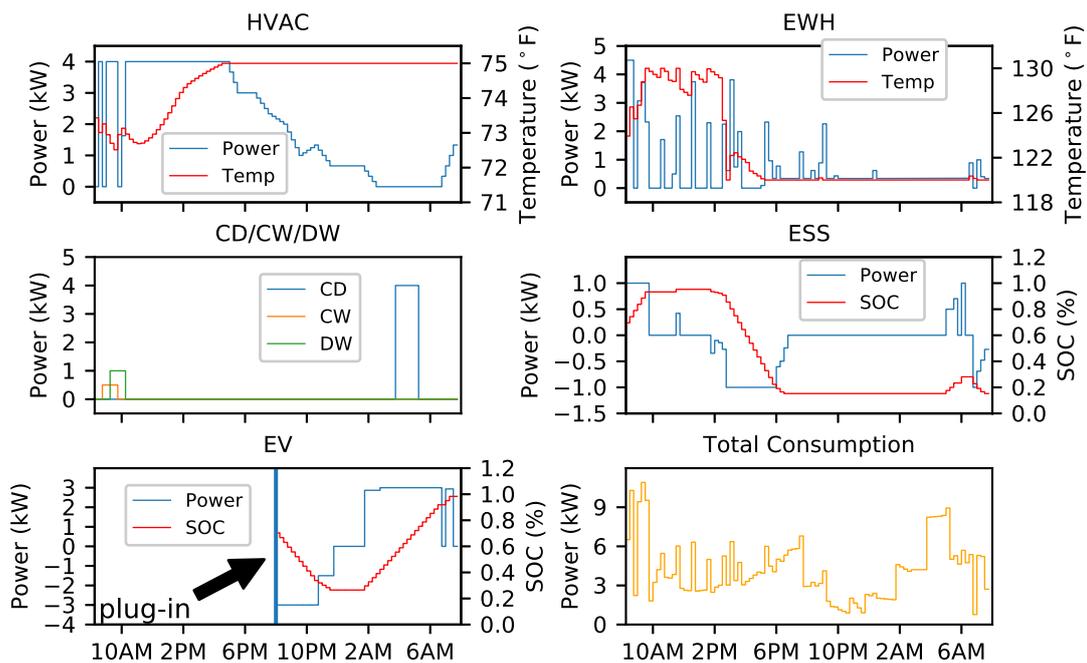

Figure 5. Optimal DR solutions of appliances under S4 based on the ADP.

4.2. Different Choices of Approximating Functions

This subsection evaluates how the choice of the approximating function affects the performance of the ADP algorithm. Except for the cubic polynomial in Equation (33) shown in Section 3.3, two additional polynomials of degree two and degree five and two RBF (radial basis function)-based approximating functions are evaluated and compared.

The two additional polynomials are designed as follows. The quadratic polynomial consists of $1, 2$ power-terms of each state variable, two-degree cross-terms and a constant term. The quintic polynomial consists of 1–5 power-terms of each state variable, $2, 3$-degree cross-terms of any two state variables and a constant term. For simplicity, denote the quadratic, cubic and quintic polynomial functions as Poly-2, Poly-3 and Poly-5, respectively.

The RBF-based approximating functions are designed based on normalized kernel functions, in which the basic functions $\phi_{i,f}(S_i)$ are defined as below,

$$\phi_{i,f}(S_i) = \frac{K(S_i, s_i^f)}{\sum_{f'=1}^{F} K(S_i, s_i^{f'})} \quad \text{for all } f \in F \tag{34}$$

where $K(S_i, s_i^f)$ is the kernel function and s_i^f is the center of the f-th kernel. The kernel function is normally a local weighting function, whose value declines as the query point goes away from its center. This enables an approximating function not only to characterize the local features of the value function in the neighborhood of every kernel center, but also to offer a proper fit for 'the middle area' through the linear combination of multiple normalized kernel functions. Clearly, the more kernel functions are used, the better the approximation that can be obtained. To show this relationship, two kernel-based approximating functions with different numbers of kernels are implemented.

For both approximating functions, the Gaussian kernels given by $K(S_i, s_i^f) = e^{-\frac{1}{2}[S_i - s_i^f]' B_f^{-1}[S_i - s_i^f]}$, which are often referred to as radial basis functions (RBFs), are applied, where B_f is the width matrix. Denote the first approximating function as RBF-1 and the second one as RBFs-2. For RBFs-1, the centers of the RBFs are arranged on a $2 \times 1 \times 2 \times 1 \times 2 \times 2 \times 2$ grid over the state space, i.e., $s_i^f = \{72, 74\} \times \{125\} \times \{0.4, 0.8\} \times \{0.6\} \times \{1, 3\} \times \{2, 4\} \times \{1, 3\}$, so there are $2^5 = 32$ RBFs in total. The width matrices are set to be $B_f = diag(2, 5, 0.4, 0.5, 2, 2, 2)$ for all $f \in F$. For RBFs-2, the centers of the RBFs are arranged on a $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2$ grid over the state space, i.e., $s_i^f = \{72, 74\} \times \{123, 127\} \times \{0.4, 0.8\} \times \{0.4, 0.8\} \times \{1, 3\} \times \{2, 4\} \times \{1, 3\}$, so there are $2^7 = 128$ RBFs in total. The width matrices are set to be $B_f = diag(2, 4, 0.4, 0.4, 2, 2, 2)$ for all $f \in F$.

Figure 6 shows the comparison using different approximating functions. For the polynomial-based approximating functions, the comparison suggests that as the degree of the polynomial goes up, the obtained total electricity cost reduces, but the CPU time increases quickly. Similar properties can be seen for the RBF-based approximating functions. This is due to the fact that a good performance relies on a good approximation, which generally contains a large number of parameters to learn, which is time consuming. In practice, a balance between the performance and the running time is needed no matter which approximating function is adopted. However, the RBF-based approximation functions require much more running time than the polynomial approximators and hence are not proper for the integrative DR problem. In addition, compared to Poly-2 and Poly-5, Poly-3 offers a good balance between the total energy cost and the CPU time according to the comparison shown in Figure 6b. Therefore, Poly-3 is an appropriate approximating function for the integrative DR problem.
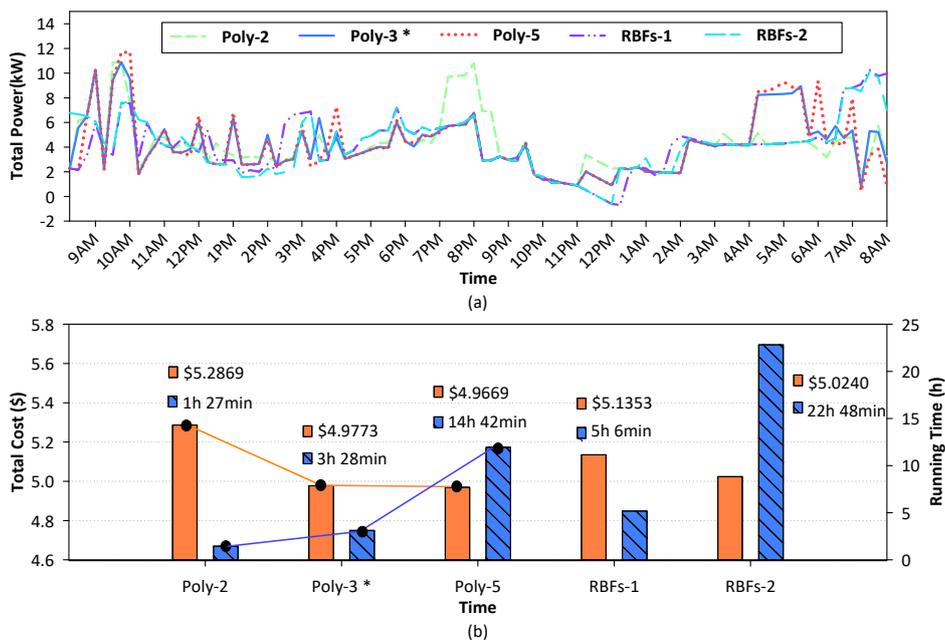

Figure 6. Comparison among different approximating functions. Poly, polynomial. (**a**) Total consumption power at each time step; (**b**) total electricity cost and running time

4.3. Comparison of Different Scenarios

Figure 7 compares the total power consumption in each time slot and the cumulative cost under different scenarios to show how the participation of different appliances affects the DR efficiency. In the figure, None stands for the scenario without DR. In this scenario, the HVAC turns ON when the indoor temperature increases above $T_{set}^{HVAC} + \Delta T^{HVAC}$ and turns off when it decreases below $T_{set}^{HVAC} - \Delta T^{HVAC}$. The operation of EWH is similar to the HVAC. The operation times of the CW, CD and DW are fixed between 5:00 p.m.–6:00 p.m., 6:00 p.m.–7:30 p.m. and 9:30 p.m.–10:30 p.m. respectively. The PV panels and the ESS are not available. The EV serves as a load and is charged immediately after arriving home.

As can be seen in the figure, under all DR scenarios, the peak load in high price hours 3 p.m.–9 p.m. and the total electricity cost are reduced compared to the scenario without DR. Specifically, the reduction becomes larger and larger as more appliances participate in DR from S1–S4. Moreover, a sharp reduction in cost happens between S2 and S3. It is obvious that the ESS and PV make a great contribution to reducing the electricity cost.
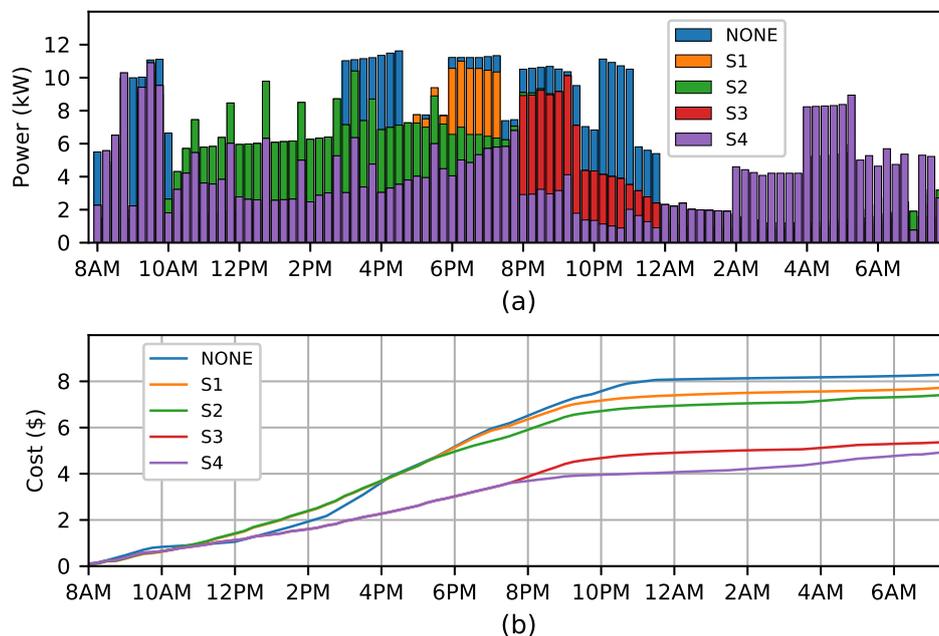

Figure 7. (**a**) Power exchange between the household and grid under different scenarios. (**b**) Cumulative cost under different scenarios.

In order to investigate how the degradation cost of the EV battery affects the V2H/V2G operation, we make a comparison for different EV battery capital cost settings of $0, $400 and $800. The simulation results are depicted in Figure 8. As can be seen, when the capital cost is $0, the SOC of the EV battery decreases to its lowest level after it arrives home, which means it deliveries the most power to the home. However, when the capital cost goes up, the SOC of the EV battery does not decrease as much. From the results, it is concluded that higher capital cost leads to less energy discharged to supply the home appliances (V2H) or to sell to the grid (V2G).

4.4. Discussion and Future Work

In developing the integrative DR strategy, we only consider from the users' perspective and assume that the bidding process has already cleared. However, it is worth mentioning that when a significant fraction of households takes the developed DR strategy and selfishly minimizes their own bills, the local demand may shift to the low price hours and form a new load peak. This certainly

would require the electric utilities to adjust their dynamic price structures by considering all kinds of loads, including residential, industry and commercial loads until a new balance between the utilities and energy consumers is achieved. This is a significant issue worth discussion from a more systemic perspective. Pioneering studies on DR pricing problems can be found in [36–40], and we refer the interesting readers to these publications. For now, we leave this problem for our future research.
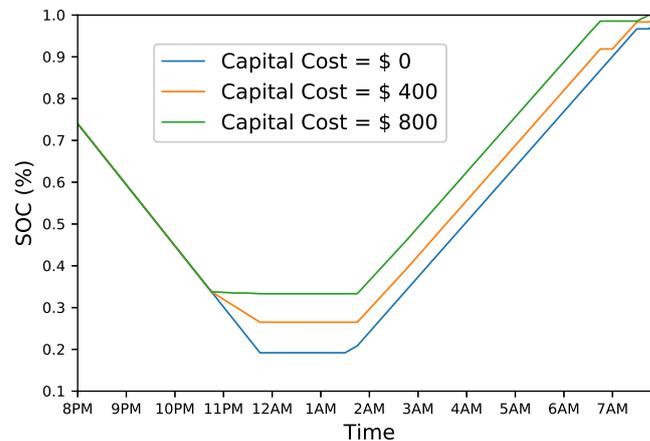
Figure 8. Comparison of the EV battery SOC and degradation cost for different capital cost settings.

5. Conclusions

This paper presents an integrative DR study for the optimal operation of home appliances, ESS and the V2G/V2H-enabled EV based on the proposed ADP algorithm. Based on the simulation results, we have the following conclusions: (1) the proposed ADP algorithm is effective and has special strength for complicated high-dimensional optimization problems; (2) more participation of smart home appliances in DR program brings more benefits to the customers via energy cost reduction and peak load shifting; (3) the ESS makes the greatest contribution to reducing the energy cost by charging from the PV or the grid and discharging to power the loads in high price hours; (4) the V2G/V2H applications of the EV can offer a more economically-efficient usage of electricity, but the degradation of the EV battery must be evaluated carefully.

Author Contributions: Hepeng Li. conceived the main idea and Peng Zeng designed the experiments; Hepeng performed the experiments; Chuanzhi Zang analyzed the data; Haibin Yu contributed analysis tools; Shuhui Li wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

P	Power (kW)
T	Temperature (°F)
E	Energy level of battery (kWh)
SOC	Battery state of charge (kWh)
Δ*T*	Allowable deviation of temperature (°F)
η	Efficiency
δ	Inertia factor of air in the house
A	thermal conductivity of the house (kW/°F)

SA	Tank surface area of EWH (ft^2)
R	Tank insulation thermal resistance of EWH (hour·ft^2·°F/BTU)
d_{water}	Density of water (8.34 lbs/gallon)
F	Hot water flow rate (gallons/hour)
C_p	Specific heat of water (1.00 BTU/(lbs·°F))
$volumn$	Capacity of the EWH tank (gallons)
s	On/off status of a shiftable appliance
L	EV battery life
DOD	Battery depth of discharge
$price$	Electricity price ($/kWh)
C	Cost ($)
V	Value function
S	State vector
a	Action vector
$\mathcal{A}$	Set of all feasible actions
π	DR control policy
θ	Regression parameters of the value function approximator
Δt	Duration of a time slot
Superscripts:	
$HVAC$	Heating, ventilation and air conditioning
EWH	Electric water heater
CW	Cloth washer
CD	Cloth dryer
DW	Dish washer
ESS	Energy storage system
EV	Electric vehicle
PV	Photovoltaics
adj	Adjustable appliances
shi	Shiftable appliances
cri	Critical appliances
(n)	Index of iteration for approximate optimistic policy iteration (AOPI) algorithm
(n, m)	Index of iteration for policy evaluation in the n-th iteration of AOPI
$*$	Optimum
Subscripts:	
i	Index of time slot
in	Indoor
out	Outdoor
set	Setting value by users
max	Maximum value
min	Minimum value
$cold$	Cold water
ch	Charging
dis	Discharging
deg	Degradation of the battery

References

1. Deng, R.; Yang, Z.; Chow, M.Y.; Chen, J. A survey on demand response in smart grids: Mathematical models and approaches. *IEEE Trans. Ind. Inform.* **2015**, *11*, 570–582.

2. Association of Home Appliance Manufactures (AHAM). Smart Grid White Paper: The Home Appliance Industrys Principles & Requirements for Achieving a Widely Accepted Smart Grid. Available online: https://www.smartgrid.gov/document/smart_grid (accessed on 20 May 2015).

3. Zhao, Z.; Lee, W.C.; Shin, Y.; Song, K.-B. An optimal power scheduling method for demand response in home energy management system. *IEEE Trans. Smart Grid* **2013**, *4*, 1391–1400.

4. Vardakas, J.; Zorba, N.; Verikoukis, C. A survey on demand response programs in smart grids: Pricing methods and optimization algorithms. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 152–178.

5. Zhang, D.; Li, S.; Sun, M.; O'Neill, Z. An Optimal and Learning–Based Demand Response and Home Energy Management System. *IEEE Trans. Smart Grid* **2016**, *7*, 1790–1801.

6. Erdinc, O.; Tascikaraoglu, A.; Paterakis, N.G.; Eren, Y.; Catalao, J.P.S. End-User Comfort Oriented Day-Ahead Planning for Responsive Residential HVAC Demand Aggregation Considering Weather Forecasts. *IEEE Trans. Smart Grid* **2017**, *8*, 362–372.

7. Pourmousavi, S.A.; Patrick, S.N.; Nehrir, M.H. Real-Time Demand Response Through Aggregate Electric Water Heaters for Load Shifting and Balancing Wind Generation. *IEEE Trans. Smart Grid* **2014**, *5*, 769–778.

8. Wei, Q.; Liu, D.; Shi, G.; Liu, Y. Multibattery Optimal Coordination Control for Home Energy Management Systems via Distributed Iterative Adaptive Dynamic Programming. *IEEE Trans. Smart Grid* **2015**, *7*, 4203–4214.

9. Althaher, S.; Mancarella, P.; Mutale, J. Automated Demand Response From Home Energy Management System Under Dynamic Pricing and Power and Comfort Constraints. *IEEE Trans. Smart Grid* **2015**, *6*, 1874–1883.

10. Anvari-Moghaddam, A.; Monsef, H.; Rahimi-Kian, A. Optimal Smart Home Energy Management Considering Energy Saving and a Comfortable Lifestyle. *IEEE Trans. Smart Grid* **2015**, *6*, 324–332.

11. Roh, H.T.; Lee, J.W. Residential Demand Response Scheduling With Multiclass Appliances in the Smart Grid. *IEEE Trans. Smart Grid* **2016**, *7*, 94–104.

12. Muratori, M.; Rizzoni, G. Residential Demand Response: Dynamic Energy Management and Time-Varying Electricity Pricing. *IEEE Trans. Power Syst.* **2016**, *31*, 1108–1117.

13. Erdinc, O.; Paterakis, N.G.; Mendes, T.D.P.; Bakirtzis, A.G.; Catalão, J.P.S. Smart Household Operation Considering Bi-Directional EV and ESS Utilization by Real-Time Pricing-Based DR. *IEEE Trans. Smart Grid* **2015**, *6*, 1281–1291.

14. Rastegar, M.; Fotuhi-Firuzabad, M. Outage Management in Residential Demand Response Programs. *IEEE Trans. Smart Grid* **2015**, *6*, 1453–1462.

15. Paterakis, N.G.; Erdinc, O.; Bakirtzis, A.G.; Catalão, J.P.S. Optimal Household Appliances Scheduling Under Day-Ahead Pricing and Load-Shaping Demand Response Strategies. *IEEE Trans. Ind. Inform.* **2015**, *6*, 1509–1519.

16. U.S. Department of Commerce. Electricity Storage in Buildings for Residential Sector Demand Response: Control Algorithms and Economic Viability Evaluation. Available online: http://dx.doi.org/10.6028/NIST.GCR.14-978 (accessed on 15 April 2016).

17. Farzin, H.; Fotuhi-Firuzabad, M.; Moeini-Aghtaie, M. A Practical Scheme to Involve Degradation Cost of Lithium–Ion Batteries in Vehicle-to-Grid Applications. *IEEE Trans. Sustain. Energy* **2016**, *7*, 1730–1738.

18. Erdinc, O. Economic impacts of small-scale own generating and storage units, and electric vehicles under different demand response strategies for smart households. *Appl. Energy* **2014**, *126*, 142–150.

19. Mohsenian-Rad, A.H.; Wong, V.W.S.; Jatskevich, J.; Schober, R.; Leon-Garcia, A. Autonomous Demand-Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid. *IEEE Trans. Smart Grid* **2010**, *1*, 320–331.

20. Powell, W.B. *Approximate Dynamic Programming*, 2nd ed.; David, N.A.C., Balding, J., Eds.; Publishing House: Westfield, NJ, USA, 2011.

21. Samadi, P.; Mohsenian-Rad, H.; Wong, V.; Schober, R. Real-time pricing for demand response based on stochastic approximation. *IEEE Trans. Smart Grid* **2014**, *5*, 789–798.

22. Simao, H.; Jeong, H.; Defourny, B.; Powell, W.; Boulanger, A.; Gagneja, A.; Wu, L.; Anderson, R. A robust solution to the load curtailment problem. *IEEE Trans. Smart Grid* **2013**, *4*, 2209–2219.

23. Wei, Q.; Liu, D.; Shi, G. A novel dual iterative Q-learning method for optimal battery management in smart residential environments. *IEEE Trans. Ind. Electron.* **2015**, *62*, 2509–2518.

24. Fuselli, D.; Angelis, F.D.; Boaro, M.; Squartini, S.; Wei, Q.; Liu, D.; Piazza, F. Action dependent heuristic dynamic programming for home energy resource scheduling. *Int. J. Electr. Power Energy Syst.* **2013**, *48* 148–160.

25. Black, J.W. Integrating Demand into the U.S. Electric Power System: Technical, Economic, and Regulatory Frameworks for Responsive Load. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2005. Available online: http://hdl.handle.net/1721.1/31168 (accessed on 5 August 2015).

26. Nehrir, M.; Jia, R.; Pierre, D.; Hammerstrom, D. Power management of aggregate electric water heater loads by voltage control. In Proceedings of the 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, USA, 24–28 June 2007; pp. 1–6.

27. Zhou, C.; Qian, K.; Allan, M.; Zhou, W. Modeling of the cost of ev battery wear due to v2g application in power systems. *IEEE Trans. Energy Convers.* **2011**, *26* 1041–1050.

28. Bellman, R. *Dynamic Programming*, 6th ed.; David, N.A.C., Balding, J., Eds.; Princeton University: Princeton, NJ, USA, 1957.

29. Davis, P.J. *Interpolation and Approximation*, 1st ed.; ser. Dover Books on Mathematics; Dover: New York, NY, USA, 2014.

30. Bertsekas, D.P.; Tsitsiklis, J.N. *Neuro-Dynamic Programming*, 1st ed.; Athena Scientific: Belmont, MA, USA, 1996.

31. Champaign-Urbana. Department of Atmospheric Sciences, University of Illinois. Available online: https://www.atmos.illinois.edu/weather/daily/index.html (accessed on 10 July 2016).

32. Iea/ecbcs Annex 42. Available online: http://www.ieaannex54.org/annex42/index.html (accessed on 10 July 2016).

33. EERE. Commercial and Residential Hourly Load Profiles for All tmy3 Locations in The United States. Available online: http://en.openei.org/datasets/?les/961/pub/ (accessed on 10 July 2016).

34. Ameren Illinois. Available online: http://www.ameren.com/account/retail-energy (accessed on 10 July 2016).

35. Yalmip. Available online: http://users.isy.liu.se/johanl/yalmip/pmwiki.php?n=Main.HomePage (accessed on 11 July 2017).

36. Yu, M.; Hong, S.H. A Real-Time Demand-Response Algorithm for Smart Grids: A Stackelberg Game Approach. *IEEE Trans. Smart Grid* **2016**, *7*, 879–888.

37. Forouzandehmehr, N.; Esmalifalak, M.; Mohsenian–Rad, H.; Han, Z. Autonomous Demand Response Using Stochastic Differential Games. *IEEE Trans. Smart Grid* **2015**, *6*, 291–300.

38. Tran, N.H.; Tran, D.H.; Ren, S.; Han, Z.; Huh, E.N.; Hong, C.S. How Geo-Distributed Data Centers Do Demand Response: A Game-Theoretic Approach. *IEEE Trans. Smart Grid* **2016**, *7*, 937–947.

39. Maharjan, S.; Zhu, Q.; Zhang, Y.; Gjessing, S.; Basar, T. Dependable Demand Response Management in the Smart Grid: A Stackelberg Game Approach. *IEEE Trans. Smart Grid* **2013**, *7*, 120–132.

40. Chai, B.; Chen, J.; Yang, Z.; Zhang, Y. Demand Response Management With Multiple Utility Companies: A Two-Level Game Approach. *IEEE Trans. Smart Grid* **2014**, *5*, 722–731.