# A Novel Iterative Adaptive Critic Design for Smart Home Energy Systems With Solar Energy

Zehua Liao, Qinglai Wei, Hongyang Li
School of Artificial Intelligence, University of Chinese Academy of Sciences
The State Key Laboratory of Management and Control for Complex Systems
Beijing, China
Email: {liaozehua2017, qinglai.wei, hongyang.li}@ia.ac.cn

2022 China Automation Congress (CAC) | 978-1-6654-6533-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/CAC57257.2022.10055827

***Abstract*—As a distributed energy storage system, smart home energy (SHE) system can be used to reduce the consumption cost of household users. Aiming at the optimal control SHE systems, a novel multi-iteration adaptive dynamic programming (MIADP) method is constructed. The present algorithm introduces three iterations, namely, exterior iteration, local iteration and interior iteration, to ensure that the iterative $Q$-function converges to the optimal, and thus the optimal energy storage control scheme of the SHE system with solar energy is obtained. Finally, the present MIADP algorithm is illustrated by the experimental results.**

***Index Terms*—Adaptive dynamic programming, smart home, adaptive critic designs, energy management, optimal control.**

## I. Introduction

SMART home energy (SHE) systems are an evolution of the present power grid structures due to the developing demand of energy, the growing charge of fossil fuels, and the improvement of new and progressive intelligent sciences [1]. Generally, a SHE system is an energy internet which is composed of energy storage equipments, power resources, and load demand. The power resources generally contain the power grids and the renewable resources which include solar energy, wind energy, and so on. The researches on power resource optimization focus on the power management of the distributed renewable resources [2]. It is challenging to construct the specific models for the SHE systems, which arouses our research interest.

Adaptive dynamic programming (ADP) is an effective technique to solve optimal control problems for nonlinear systems [3], and the kernel technology is to employ a critic module to evaluate the performance and simultaneously direct the actor to search a better control law. ADP has been employed for solving the optimal energy management of the SHE systems [4]–[6]. Inspired by [6], we consider developing a novel method to achieve the optimal control of SHE systems.

In this paper, a new algorithm called MIADP is developed for obtaining the optimal control of SHE with solar renewable energy. The numerical results and the comparisons with the methods proposed by [2], [4], [7] will illustrate the effectiveness and superiority of the present MIADP method.

This work was supported in part by the National Key R&D Program of China under Grants 2021YFE0206100; in part by the Science and Technology Development Fund, Macau SAR under Grants 0060/2021/A2 and 0015/2020/AMJ; in part by National Defense Basic Scientific Research Program JCKY2019203C029; in part by the National Natural Science Foundation of China under Grant 62073321.

## II. Problem Formulations

### A. SHE System Description

The studied SHE system shown in Fig. 1 can be seen as an energy internet composed of power resources, load demand, and energy storage equipments.

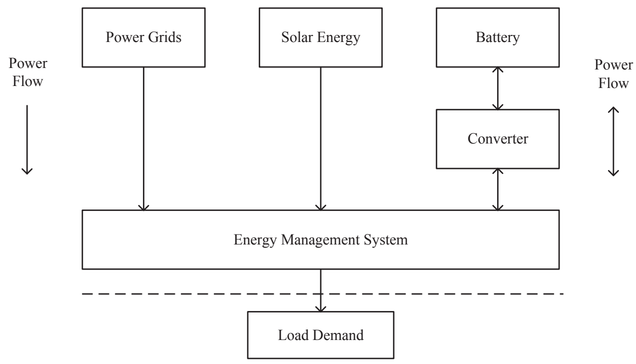


Fig. 1. SHE System Structure.

The load balance is

$$E_{l,\kappa} = E_{bl,\kappa} + E_{gl,\kappa} + E_{rl,\kappa}, \tag{1}$$

where $\kappa$ is the time index, "$l$", "$b$", "$g$" and "$r$" denote "load", "battery", "grid" and "solar renewable energy", respectively. We use $E_{bl,\kappa}$ (kW) to denote the energy from the battery to the load, and other notations have similar meanings. Besides, we also have $E_{g,\kappa} = E_{gl,\kappa} + E_{gb,\kappa}$ and $E_{r,\kappa} = E_{rl,\kappa} + E_{rb,\kappa}$. The battery model is

$$\begin{aligned}\Xi_{b,\kappa+1} =& \Xi_{b,\kappa} - E_{bl,\kappa} \times \eta(E_{bl,\kappa}) \\ &+ \big(E_{rb,\kappa} + E_{gb,\kappa}\big) \times \eta\big(E_{rb,\kappa} + E_{gb,\kappa}\big),\end{aligned} \tag{2}$$

where $\Xi_{b,\kappa}$ (kWh) is the battery energy, and the efficiency of battery is defined as $\eta\big(E_{rb,\kappa} + E_{gb,\kappa}\big) = 0.898 - 0.173 \left|E_{rb,\kappa} + E_{gb,\kappa}\right| / E_{\text{rate}}$ according to [4], where $E_{\text{rate}}$ is the rated power.

Since the solar energy is free, it is first needed to meet the load power supply, then be charged for energy storage equipment. Define the new renewable source and load as

$$\mathcal{E}_{r,\kappa} = \begin{cases} E_{r,\kappa} - E_{l,\kappa}, & E_{l,\kappa} - E_{r,\kappa} \geq 0 \\ 0, & E_{l,\kappa} - E_{r,\kappa} < 0 \end{cases} \tag{3}$$

and

$$\mathcal{E}_{l,\kappa} = \begin{cases} E_{l,\kappa} - E_{r,\kappa}, & E_{l,\kappa} - E_{r,\kappa} \geq 0 \\ 0, & E_{l,\kappa} - E_{r,\kappa} < 0 \end{cases} . \tag{4}$$

The new load balance will be

$$\mathcal{E}_{l,\kappa} = E_{bl,\kappa} - E_{gb,\kappa} + E_{g,\kappa}. \tag{5}$$

where $\mathcal{E}_{rb,\kappa}$ satisfies $\mathcal{E}_{rb,\kappa} = \mathcal{E}_{r,\kappa}$.

### *B. Optimization Objectives*

Let $C_\kappa$ be the electricity rate. Based on (2)–(5), the performance index function of the SHE is

$$\sum_{t=0}^{\infty} \gamma^t \Big( \beta_1 (C_\kappa E_{g,\kappa})^2 + \beta_2 (\Xi_{b,\kappa} - \Xi_b^o)^2 + \beta_3 (E_{bl,\kappa} - E_{gb,\kappa})^2 \Big), \tag{6}$$

where $0 < \gamma < 1$, $\beta_1$, $\beta_2$ and $\beta_3$ are positive constants, $\Xi_b^o$ is the middle of battery storage limit.

Let $\varkappa_{1,\kappa} = E_{g,\kappa}$, $\varkappa_{2,\kappa} = \Xi_{b,\kappa} - \Xi_b^o$, $\varkappa_\kappa = [\varkappa_{1,\kappa}, \varkappa_{2,\kappa}]^T$. Let $u_\kappa = E_{bl,\kappa} - E_{gb,\kappa}$. The discrete SHE system function is defined as

$$\begin{aligned} \varkappa_{\kappa+1} &= S(\varkappa_\kappa, u_\kappa, t) \\ &= \begin{pmatrix} \mathcal{E}_{l,\kappa} - u_\kappa \\ \varkappa_{2,\kappa} - (u_\kappa - \mathcal{E}_{rb,\kappa})\eta(u_\kappa - \mathcal{E}_{rb,\kappa}) \end{pmatrix}. \end{aligned} \tag{7}$$

According to (6), defined the utility function as

$$U(\varkappa_\kappa, u_\kappa, t) = \beta_1 (C_\kappa \varkappa_{1,\kappa})^2 + \beta_2 {\varkappa_{2,\kappa}}^2 + \beta_3 {u_\kappa}^2. \tag{8}$$

Based on the utility function, the optimal performance index function satisfies the following Bellman equation:

$$J^*(\varkappa_\kappa, t) = \inf_{\underline{u}_\kappa} \left\{ U(\varkappa_\kappa, u_\kappa, t) + \gamma J(\varkappa_{\kappa+1}, \kappa+1) \right\}. \tag{9}$$

## III. Iterative Adaptive Critic Designs for SHE System

Assume that the energy data and electricity price of the SHE system change periodically, with a period of $\theta = 24$ hours. Redefine utility function as

$$\Psi(\varkappa_\tau, \mathfrak{U}_\tau) = \sum_{\epsilon=0}^{\theta-1} \gamma^\theta U(\varkappa_{\tau+\epsilon}, u_{\tau+\epsilon}, \epsilon), \tag{10}$$

for $\tau \in \{0, \theta, 2\theta, \ldots\}$. Rewrite (9) as

$$Q^*(\varkappa_\tau, \mathfrak{A}_\tau) = \Psi(\varkappa_\tau, \mathfrak{U}_\tau) + \rho \min_{\mathfrak{U}_{\tau+\theta}} Q^*(\varkappa_{\tau+\theta}, \mathfrak{U}_{\tau+\theta}), \tag{11}$$

where $\rho = \gamma^\theta$. The corresponding optimal control $\mathcal{U}^*(\varkappa_\tau) = \arg\min_{\mathfrak{U}_\tau} Q^*(\varkappa_\tau, \mathfrak{U}_\tau)$.

Based on the preparations above, the MIADP algorithm is developed with three iteration processes. The first iteration is the exterior iteration, whose objective is to achieve the optimum by updating the iterative $Q$ function. Let $\mathcal{U}_0(\varkappa_\tau)$ denote the initial control law, the initial value function is compute as

$$Q_0(\varkappa_\tau, \mathfrak{U}_\tau) = \Psi(\varkappa_\tau, \mathfrak{U}_\tau) + \rho Q_0(\varkappa_{\tau+\theta}, \mathcal{U}_0(\varkappa_{\tau+\theta})). \tag{12}$$

Then, for $i = 1, 2, \ldots$, the exterior iteration is proceeded between

$$\mathcal{U}_i(\varkappa_\tau) = \arg\min_{\mathfrak{U}_\tau} Q_{i-1}(\varkappa_\tau, \mathfrak{U}_\tau) \tag{13}$$

and

$$Q_i(\varkappa_\tau, \mathfrak{U}_\tau) = \Psi(\varkappa_\tau, \mathfrak{U}_\tau) + \rho Q_i(\varkappa_{\tau+\theta}, \mathcal{U}_i(\varkappa_{\tau+\theta})). \tag{14}$$

For the second iteration process, let $\mathcal{L} = \{\mathcal{L}_1, \mathcal{L}_2, \ldots\}$ be a set of positive integers sequence. Define $l_1 = 0, 1, \ldots, \mathcal{L}_1 - 1$. For $i = 1$, there is

$$Q_{1,l_1+1}(\varkappa_\tau, \mathfrak{U}_\tau) = \Psi(\varkappa_\tau, \mathfrak{U}_\tau) + \rho Q_{1,l_1}(\varkappa_{\tau+\theta}, \mathcal{U}_1(\varkappa_{\tau+\theta})), \tag{15}$$

where $Q_{1,0}(\varkappa_\tau, \mathfrak{U}_\tau) = Q_0(\varkappa_\tau, \mathfrak{U}_\tau)$. For $i = 2, 3, \ldots$, define

$$Q_{i,0}(\varkappa_\tau, \mathfrak{U}_\tau) = Q_{i-1}(\varkappa_\tau, \mathfrak{U}_\tau) = Q_{i-1,\mathcal{L}_{i-1}}(\varkappa_\tau, \mathfrak{U}_\tau). \tag{16}$$

Obtain $\mathcal{U}_i(\varkappa_\tau)$ by (13), we have

$$Q_{i,l_i+1}(\varkappa_\tau, \mathfrak{U}_\tau) = \Psi(\varkappa_\tau, \mathfrak{U}_\tau) + \rho Q_{i,l_i}(\varkappa_{\tau+\theta}, \mathcal{U}_i(\varkappa_{\tau+\theta})). \tag{17}$$

The third iteration is the interior iteration to solve (13). Define $j_i = 0, 1, \ldots, \theta - 1$. Define

$$Q_0^0(\varkappa_\tau, u_\tau) = Q_0(\varkappa_\tau, \mathcal{U}_0(\varkappa_\tau)). \tag{18}$$

The control law corresponding to (18) is computed as

$$u_0^0(\varkappa_\tau) = \arg\min_{u_\tau} Q_0^0(\varkappa_\tau, u_\tau). \tag{19}$$

Then, for $i = 0$ and $j_i = 1, 2, \ldots, \theta$, the internal iteration is proceeded between

$$Q_0^{j_i}(\varkappa_\tau, u_\tau) = \mathcal{U}(\varkappa_\tau, u_\tau, j_i) + \gamma Q_0^{j_i-1}(\varkappa_{\tau+1}, u_0^{j_i-1}(\varkappa_{\tau+1})) \tag{20}$$

and

$$u_0^{j_i}(\varkappa_\tau) = \arg\min_{u_\tau} Q_0^{j_i}(\varkappa_\tau, u_\tau), \tag{21}$$

where we let

$$\varkappa_{\tau+1} = \begin{pmatrix} \mathcal{E}_{l,\theta-1-j_i} - u_\tau \\ \varkappa_{2,\tau} - (u_\tau - \mathcal{E}_{rb,\tau})\eta(u_\tau - \mathcal{E}_{rb,\tau}) \end{pmatrix}. \tag{22}$$

and define the utility function as

$$\Pi(\varkappa_\tau, u_\tau, j_i) = \beta_1 \big( C_{(\theta-1-j_i)} \varkappa_{1,\tau} \big)^2 + \beta_2 {\varkappa_{2,\tau}}^2 + \beta_3 {u_{,\tau}}^2. \tag{23}$$

For $i = 1, 2, \ldots$, we let $Q_i^0(\varkappa_\tau, u_\tau) = Q_{i-1}^\theta(\varkappa_\tau, u_\tau)$. The corresponding control law is

$$u_i^0(\varkappa_\tau) = \arg\min_{u_\tau} Q_i^0(\varkappa_\tau, u_\tau). \tag{24}$$

Then, for $j_i = 1, 2, \ldots, \theta$, the interior iteration is proceeded between

$$Q_i^{j_i}(\varkappa_\tau, u_\tau) = \Pi(\varkappa_\tau, u_\tau, j_i) + \gamma Q_i^{j_i-1}(\varkappa_{\tau+1}, u_i^{j_i-1}(\varkappa_{\tau+1})) \tag{25}$$

and

$$u_i^{j_i}(\varkappa_\tau) = \arg\min_{u_\tau} Q_i^{j_i}(\varkappa_\tau, u_\tau). \tag{26}$$

Hence, $\mathcal{U}_i(\varkappa_\tau)$ can be obtained by

$$\mathcal{U}_i(\varkappa_\tau) = \left\{u_i^{\theta-1}(\varkappa_\tau), u_i^{\theta-2}(\varkappa_\tau), \ldots, u_i^0(\varkappa_\tau)\right\}. \tag{27}$$

## IV. Numerical Example

### A. Preparation

Settings before simulations.

- The simulation period is one week (168 h).
- Let $\Xi_{b,\kappa} \in [20, 80]$ kWh for safe use, and $\Xi_{b,0} = 50$ kWh.
- $E_{\text{rate}} = 16$ kW.
- $\beta_1 = 1$, $\beta_2 = 0.4$ and $\beta_3 = 0.2$.
- $\gamma = 0.99$.

The periodic data of load, electricity price and the solar energy are shown in Fig. 2 to Fig. 4, respectively, which are taken from [2], [4].

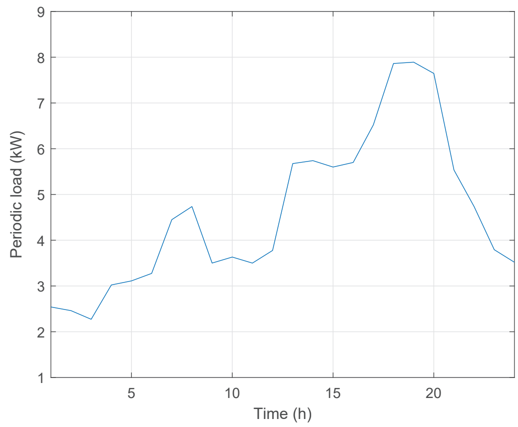


Fig. 2. Load demand

### B. Results and Comparisons

Implement MIADP for $i = 15$ iterations. The convergence curve in Fig. 5 shows that the $Q$-function converges after 10 exterior iterations. The corresponding optimal management scheme of the SHE system is displayed in Fig. 6. We can derive that, when the electricity rate is low, the battery charged from the grid. When the electricity rate is high, the battery discharge to meet the load, which reduces the power cost. Besides, the solar energy is first meet the load, and the remaining energy is charged for the battery.

To show the superiority of the MIADP algorithm, we employed the PSO method [2], the TBQL method [4] and

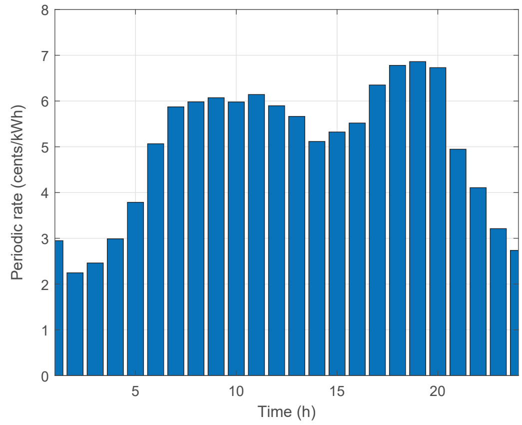


Fig. 3. Electricity rate

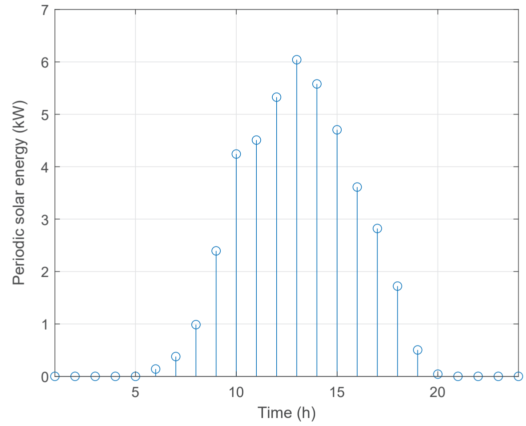


Fig. 4. Solar energy

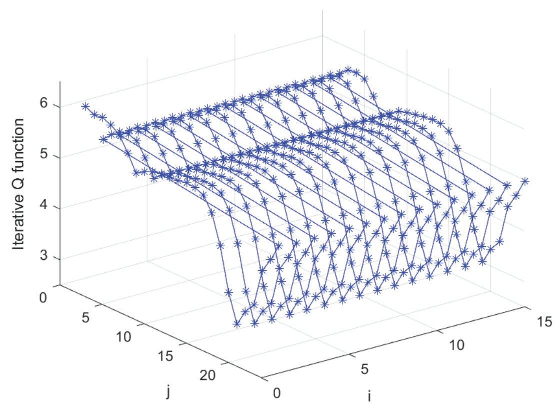


Fig. 5. Convergence trajectory of MIADP algorithm

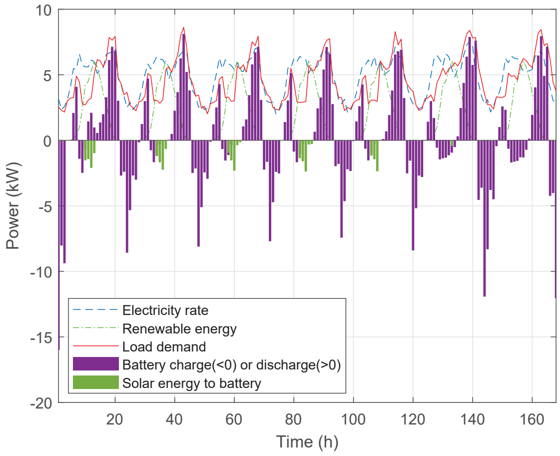


Fig. 6. Optimal management

the Dual QL method [7] for comparisons. Table I shows the comparison results, which prove that implementing the MIADP method can obtain better results. The effectiveness and superiority of the MIADP are illustrated.

TABLE I
COMPARISON RESULTS

| | Original | PSO | TBQL | Dual QL | MIADP |
|---|---|---|---|---|---|
| Cost (cents) | 2791.46 | 1795.96 | 1738.92 | 1717.32 | 1713.32 |
| Saving rate | | 35.66% | 37.71% | 38.48% | 38.62% |
| Iteration No. | | | 200 | 30 | 10 |

Next, new simulation data is adopted. The two new sets of solar energy data are displayed in Fig. 7 and Fig. 8, respectively. The corresponding optimal management results are shown in Fig. 9 and Fig. 10, respectively. When the solar energy is small, it will be fully used to supply the load demand. When the solar energy is large, a large amount of spilled energy will try to fill the battery, but energy waste may occur.

## V. CONCLUSION

In this paper, a multi-iteration ADP method is presented, which can solve the optimal energy management problem of SHE system with solar energy. Numerical results show that the MIADP method can minimize the cost of the SHE system and has better performance than other commonly used energy control algorithms. In the future research, we will consider how to deal with the possible surplus renewable energy, prevent the generation of energy waste, and obtain better economic benefits.

## REFERENCES

[1] M. Yousefi, A. Hajizadeh, M. N. Soltani, and B. Hredzak, "Predictive home energy management system with photovoltaic array, heat pump, and plug-in electric vehicle," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 430–440, Jan. 2021.

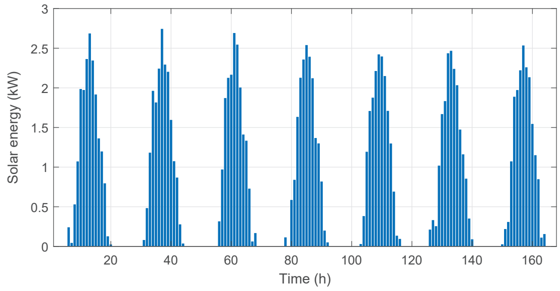


Fig. 7. New solar energy data 1

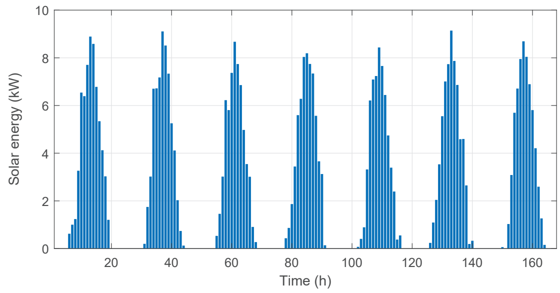


Fig. 8. New solar energy data 2

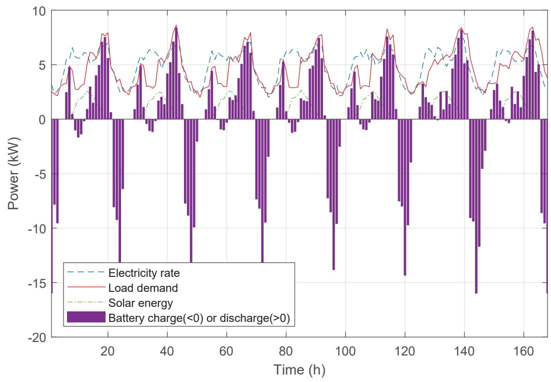


Fig. 9. Optimal management with data 1

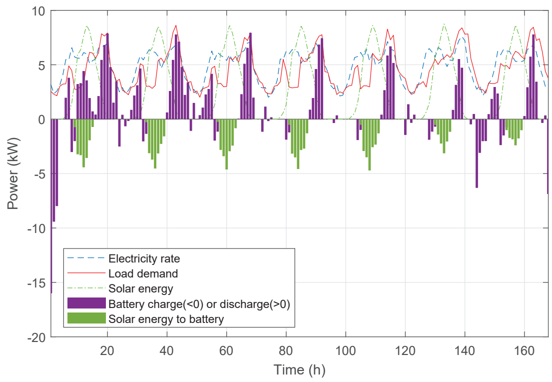


Fig. 10. Optimal management with data 2

[2] D. Fuselli, F. D. Angelis, M. Boaro, D. Liu, Q. Wei, S. Squartini, and F. Piazza, "Action dependent heuristic dynamic programming for home energy resource scheduling," *International Journal of Electrical Power and Energy Systems*, vol. 48, pp. 148–160, Jun. 2013.

[3] P. J. Werbos, "Advanced forecasting methods for global risis warning and models of intelligence," *General Systems Yearbook*, vol. 22, pp. 25–38, 1977.

[4] T. Huang and D. Liu, "A self-learning scheme for residential energy system control and management," *Neural Computing and Applications*, vol. 22, no. 2, pp. 259–269, Feb. 2013.

[5] L. Dong, Y. Tang, H. He, and C. Sun, "An event-triggered approach for load frequency control with supplementary ADP," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 581–589, Jan. 2017.

[6] Q. Wei, Z. Liao, and G. Shi, "Generalized actor-critic learning optimal control in smart home energy management," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 6614–6623, Oct. 2021.

[7] Q. Wei, D. Liu, and G. Shi, "A novel dual iterative $Q$-learning method for optimal battery management in smart residential environments," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2509–2518, Apr. 2015.