Journal of Power Sources 196 (2011) 3280–3289

Contents lists available at ScienceDirect

Journal of Power Sources

journal homepage: www.elsevier.com/locate/jpowsour

Energy management of a fuel cell/ultracapacitor hybrid power system using an adaptive optimal-control method

Wei-Song Lin*, Chen-Hong Zheng

Department of Electrical Engineering, National Taiwan University, Taiwan

ARTICLE INFO

Article history:
Received 13 July 2010
Accepted 8 November 2010
Available online 26 November 2010

Keywords:
Fuel cell
Hybrid power system
Energy management
Adaptive optimal control

ABSTRACT

Energy management of a fuel cell/ultracapacitor hybrid power system aims to optimize energy efficiency while satisfying the operational constraints. The current challenges include ensuring that the non-linear dynamics and energy management of a hybrid power system are consistent with state and input constraints imposed by operational limitations. This paper formulates the requirements for energy management of the hybrid power system as a constrained optimal-control problem, and then transforms the problem into an unconstrained form using the penalty-function method. Radial-basis-function networks are organized in an adaptive optimal-control algorithm to synthesize an optimal strategy for energy management. The obtained optimal strategy was verified in an electric vehicle powered by combining a fuel-cell system and an ultracapacitor bank. Driving-cycle tests were conducted to investigate the fuel consumption, fuel-cell peak power, and instantaneous rate of change in fuel-cell power. The results show that the energy efficiency of the electric vehicle is significantly improved relative to that without using the optimal strategy.

1. Introduction

A fuel-cell hybrid power system (FCHPS) for an electric vehicle augments the fuel cell with a reversible energy storage system (ESS) so that the overall system can cope with the power demands of the vehicle. The ESS can be implemented with either an ultracapacitor bank or a rechargeable battery [1–3]—this work considers the ultracapacitor-based ESS. The chief merit of this technology is that the power-capacity rating of the fuel-cell system (FCS) is required to meet the average demand only, rather than the peak demand. This makes the FCHPS more cost-effective and energy-efficient than using the fuel cell alone in powering the vehicle. Secondly, rapid load variations may induce oxygen starvation and thereby cause permanent damage to the proton-exchange membranes of the fuel cell. In contrast, the ultracapacitor exhibits superior performance in providing peak power, despite its significantly low energy density. Combining an FCS and an ultracapacitor bank can provide a power system with both high power and energy densities. Thirdly, the FCS generates electric power directly from hydrogen, but a reverse power flow is impossible; the ultracapacitor bank therefore provides a reservoir for the regenerative use of electricity.

As shown in Fig. 1, the FCHPS uses DC/DC converters to interface various types of power devices. This may be accomplished using a unidirectional boost converter that interfaces with the FCS and the DC bus, and protects the FCS from damage by reverse current. Alternatively, the ESS may employ a bidirectional buck-boost converter to allow bidirectional power flow, which allows the ESS to not only deliver the shared peak power but also capture the electricity for regenerative utilization. The FCHPS employs an energy-management strategy (EMS) to achieve an optimal power split between distinct power sources that improves the energy utilization. Based on a model-predictive control methodology, Vahidi et al. [4] developed a current-management strategy to avoid the problems of oxygen starvation, air-compressor surge and choke in FCSs. Chen et al. [5] used multiple-model predictive control to optimize the power usage and the control of oxygen. Zhu et al. [6] adopted a cluster-weighted modelling algorithm to identify load transients and determine the optimal power split between the FCS and ESS. A transient-load recognition technique based on wavelet-transform algorithms for hybrid energy sources (including a fuel cell, battery and ultracapacitor) was proposed in [7]. Jiang [8] investigated using an agent-based power-sharing method for implementing a distributed control scheme when combining multiple power sources. An EMS based on fuzzy logic was

Abbreviations: FCHPS, fuel-cell hybrid power system; ESS, energy storage system; FCS, fuel-cell system; EMS, energy-management strategy; AOC, adaptive optimal control; RBF, radial basis function; SoC, state of charge; VCC, voltage-and-current controller.

* Corresponding author at: Department of Electrical Engineering, National Taiwan University, EE, No. 1, Sec. 4, Roosevelt Road, Taipei City, Taiwan.
Tel.: +886 2 33663638; fax: +886 2 23638247.
E-mail addresses: weisong@cc.ee.ntu.edu.tw, d94921006@ntu.edu.tw (W.-S. Lin), wmycode@gmail.com (C.-H. Zheng).

0378-7753/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpowsour.2010.11.127

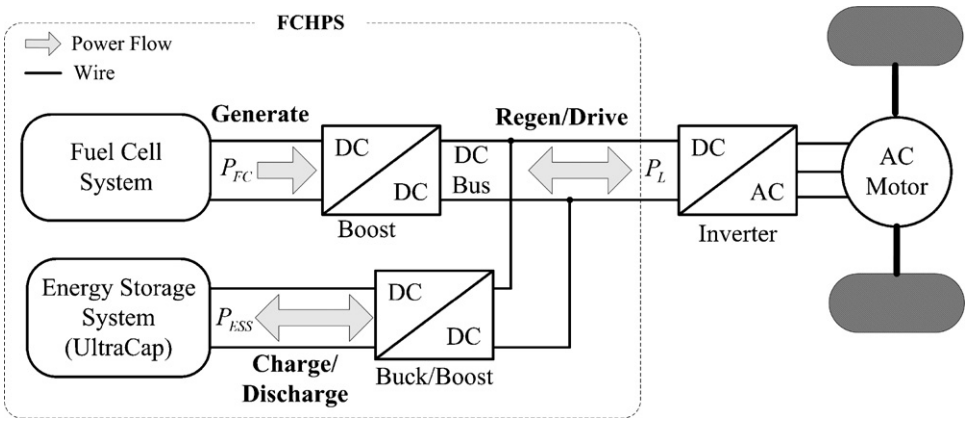

Fig. 1. Architecture of the FCHPS.

studied in [9–12]. Moreover, an EMS may be derived with the help of optimization theory. Based on the equivalent consumption-minimization strategy, Paganelli et al. [13] and Rodatz et al. [14] presented a local optimal scheme that evaluated the cost function based on the hydrogen consumption and the equivalent fuel consumption of the ESS. Feroldi et al. [15] formulated an EMS as a local optimization problem subject to a set of operational constraints, and used a constrained non-linear programming method to obtain a solution. Delprat et al. [16] and Bernard et al. [17] obtained the EMS by the forward iteration of a state equation, a costate equation and a stationary equation while assuming that the initial costate vector is available. This approach provides a non-causal solution because the entire driving cycle must be known before obtaining the initial costate vector, and the convergence of this algorithm is sensitive to parameter errors.

In the present work, the requirements for the EMS are formulated as an optimal-control problem subject to a set of state and input constraints imposed by operational limitations. The penalty-function method [18] is then used to transform the constrained optimal-control problem into an unconstrained form. An adaptive optimal-control (AOC) algorithm is subsequently developed and customized for synthesizing an optimal EMS. The AOC algorithm is deduced from the minimum principle of optimal control, rather than by manipulating the Bellman equation, as in adaptive dynamic programming [19–21]. Radial-basis-function (RBF) networks are employed when constructing the AOC algorithm. The obtained optimal EMS was evaluated in the application of an electric vehicle powered by an FCHPS. The results from driving-cycle tests demonstrate the effectiveness of the optimal EMS. Here we focus on formulating a constrained optimal-control problem for energy management of an FCHPS and on developing an AOC methodology for synthesizing an optimal EMS through reinforcement learning.

This paper is organized as follows. Section 2 derives a model for the FCHPS, and the requirements of its EMS are formulated as a constrained optimal-control problem. Section 3 develops an AOC algorithm that is capable of synthesizing an optimal EMS by learning training data. Section 4 investigates the obtained optimal EMS with driving-cycle tests of an electric vehicle powered by an FCHPS. Finally, conclusions are drawn in Section 5.

2. Formulation of the energy-management problem

2.1. Model for the FCHPS

A mathematical model for the FCS can be deduced from physical laws and the operating conditions. If the gas pressure, ambient temperature and humidity are all well-regulated inside the FCS, its characteristic curve of efficiency versus net power would exhibit

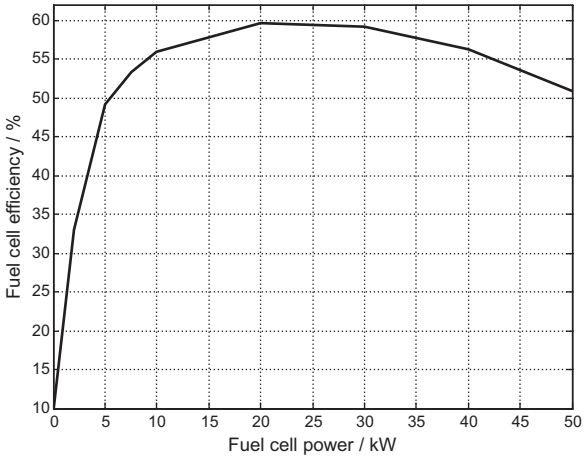

Fig. 2. Efficiency of a 50-kW fuel cell.

the result shown in Fig. 2. The power efficiency is low when the FCS operates in a low-power mode since the peripheral systems of the FCS consume an amount of power in maintaining the functionality of the overall system. In addition, the power efficiency reduces in a high-power mode due to the physical nature of fuel-cell stacks. Fig. 2 shows the overall efficiency of a 50-kW fuel-cell module as calculated using ADVISOR software [22].

However, knowledge of the overall efficiency is not sufficient for assessing the performance of an FCS, since the low efficiency in the low-power mode (Fig. 3) is not significant due to the low absolute power loss. Therefore, we gauged the performance of the FCHPS by deriving the hydrogen consumption from the efficiency of the fuel cell [11] as (see Fig. 3):

$$\dot{m}_{H_2} = \frac{P_{FC}}{LHV \cdot \eta_{FC}(P_{FC})} \times 105\%, \quad (1)$$

where $\dot{m}_{H_2}$ is the hydrogen-fuel consumption rate ($g\,s^{-1}$), LHV = 12,000 kJ g^{-1} is the lower heating value of hydrogen and the additional 5% allows for the assumed loss of hydrogen due to the FCS purging mechanism. A performance model of the FCS is obtained by fitting the hydrogen-consumption curve in Fig. 3 with the third-order polynomial function.

$$\dot{m}_{fit}(P_{FC}) = a_3 P_{FC}^3 + a_2 \cdot P_{FC}^2 + a_1 \cdot P_{FC} + a_0. \quad (2)$$

Table 1 lists the fitting parameters in Eq. (2) with $\dot{m}_{fit}$ in units of $g\,s^{-1}$ and P_{FC} in units of 100 kW (this unit is adopted for power to ensure that the polynomial fitting function does not contain extremely large or small coefficients). An appropriate parameter set can facilitate the convergence of the optimal solution search.

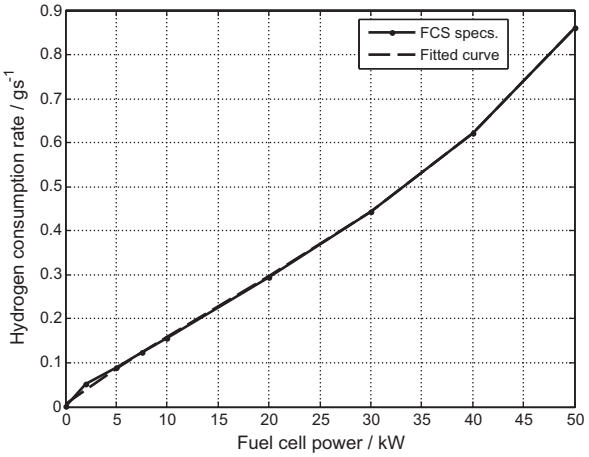

Fig. 3. Hydrogen consumption of the 50-kW fuel cell.

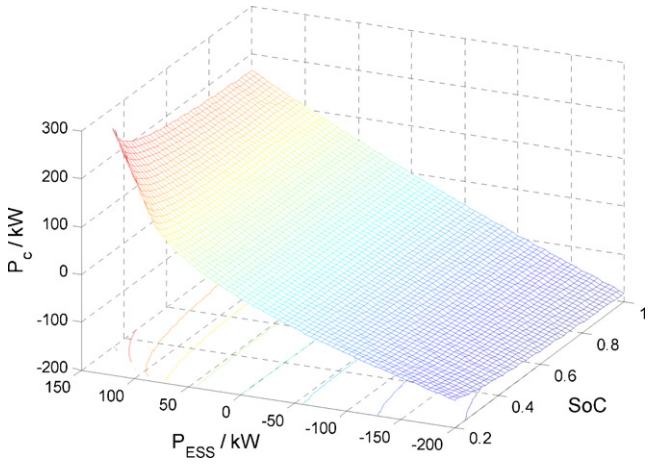

Fig. 5. Charge–discharge contour plot of an ultracapacitor bank.

Table 1
Curve-fitting parameters for a performance model.

Parameter	a_3	a_2	a_1	a_0
Value	4.1953	−2.0662	1.6944	0.0066

The ultracapacitor behaves similarly to a conventional capacitor. Previous studies have shown that an ultracapacitor can be modelled as an equivalent circuit with constant parameters [3,23–25]. Fig. 4 shows the equivalent circuit model used in this study. According to this model, the energy (E) stored in the ultracapacitor is linearly proportional to the square of the capacitor voltage. The maximum energy (E_{max}) that can be stored in an ultracapacitor is limited by its rated voltage (v_{max}) according to

$$E = \frac{1}{2}C_{uc}v_c^2 = \frac{1}{2}C_{uc}\left(\frac{P_c}{i_c}\right)^2, \tag{3}$$

$$E_{max} = \frac{1}{2}C_{uc}v_{max}^2. \tag{4}$$

The state of charge (SoC) of an ultracapacitor is defined as the ratio of E to E_{max}:

$$\text{SoC} \equiv \frac{E}{E_{max}}. \tag{5}$$

As shown in Fig. 4, P_{ESS} denotes the power flow at the terminal nodes of the ultracapacitor, and P_c denotes the power passing through internal capacitor C_{uc}. The convention used here is for discharging power and charging power to have positive and negative values, respectively. The parasitic resistance (R_{uc}) results in significant power loss during charging and discharging, and leads to a charge–discharge contour plot that takes the form

$$P_c(\text{SoC},P_{ESS}) = \left(1 - \sqrt{1 - \frac{2R_{uc}C_{uc}}{\text{SoC}\cdot E_{max}}P_{ESS}}\right)\cdot\frac{\text{SoC}\cdot E_{max}}{R_{uc}C_{uc}}. \tag{6}$$

Fig. 5 shows the deduced charge–discharge contour plot of an ultracapacitor. The dynamics of an ultracapacitor can be modelled by a first-order difference equation:

$$\text{SoC}(k+1) = \text{SoC}(k) - P_c(\text{SoC}(k), P_{ESS}(k))\cdot\frac{\Delta T}{E_{max}}, \tag{7}$$

where k (with $k = 0, 1, 2, \ldots$) is the index of the time steps and ΔT is the sampling period.

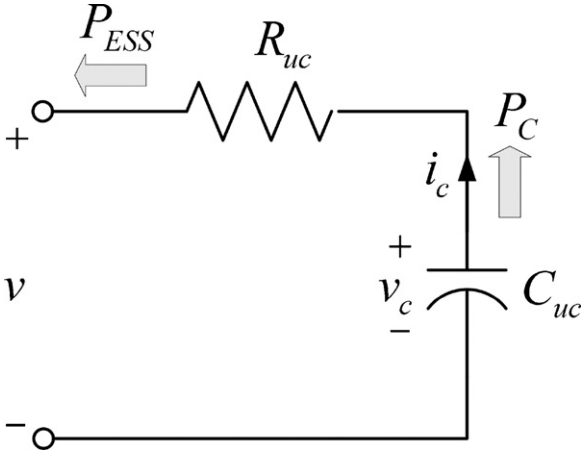

Fig. 4. Model of an ultracapacitor as an RC equivalent circuit.

2.2. Objective of energy management

Manipulating an FCHPS involves controlling the DC/DC converters and managing how power is split among distributed power sources. The converter-control loop typically responds 100–1000 times faster than the energy-management-strategy loop. Therefore, as shown in Fig. 6, the entire EMS of the FCHPS can be divided

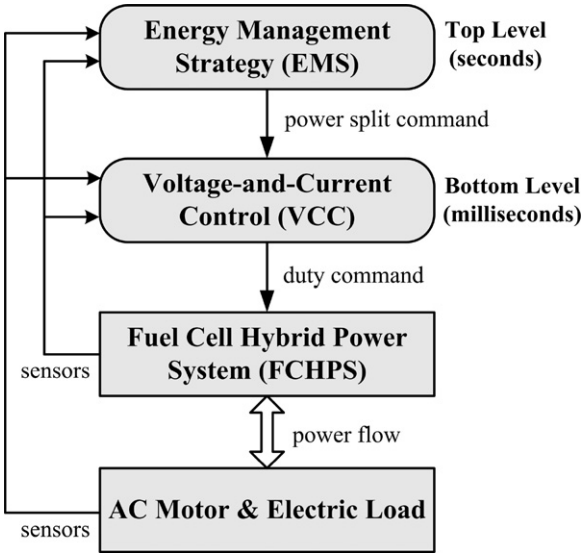

Fig. 6. Levels of the EMS and VCC in an FCHPS.

into a bottom level and a top level associated with the fast- and slow-responding loops, respectively. The bottom level is responsible for the stable control of voltage and current for the DC/DC converters; at this level, the voltage-and-current controller (VCC) acts as an interface between the top level and the physical components of the FCHPS. The top level implements the EMS that directs the FCHPS to minimize the cost function.

Methods used to control DC/DC converters are described in [26–28]. In the present study it was assumed that the converter-control loop performs so effectively that the steady-state errors and transient behaviours of the DC/DC converters can be neglected. Thus, the transfer function of the VCC was approximated as a constant value of 1, allowing the problem of energy management of the FCHPS to be formulated as a constrained optimal-control problem. Using the polynomial model of Eq. (2) of the hydrogen consumption, a cost function for assessing the performance of the FCHPS can be written as

$$J' = \Phi(K,\mathrm{SoC}(K)) + \sum_{k=0}^{K-1} \dot{m}_{\mathrm{H_2}}(P_{\mathrm{FC}}(k)) \cdot \Delta T, \tag{8}$$

where K denotes the final step of the test cycle, and

$$\Phi(K,\mathrm{SoC}(K)) = (\mathrm{SoC}(K) - \mathrm{SoC}(0))^T P_K (\mathrm{SoC}(K) - \mathrm{SoC}(0)), \tag{9}$$

which minimizes the difference between the final SoC [SoC(K)] at the end of the driving cycle and the initial SoC [SoC(0)]. This restriction ensures that the vehicle is on average powered by the FCS. This cost function accounts for the total hydrogen consumption during an entire driving cycle. The goal of energy management is to minimize the cost function of Eq. (8), subject to the constraints imposed by the ultracapacitor dynamics given by Eq. (7) and the operational limitations of the FCS and the ESS.

The principle of energy conservation means that the following equality constraint holds for this power system:

$$P_{\mathrm{FC}}(k)\eta_1 + P_{\mathrm{ESS}}(k)\eta_2 = P_L(k), \tag{10}$$

where η_1 and η_2 are the efficiencies of DC/DC converters. In real-world applications, the operating ranges of the FCS and the ESS should be appropriately restricted so as to protect the high-power equipment from damage. The lower power limitation ($P_{\mathrm{FC_{min}}} > 0$), is applicable in the FCS because it cannot store electrical energy. The FCS may be damaged permanently or it may degrade rapidly if any reverse current occurs, and hence $P_{\mathrm{FC_{min}}}$ must be sufficiently greater than 0 to avoid violating this constraint. When the FCS is operating at a high power, the maximum power that can be drawn is limited to the rated value, $P_{\mathrm{FC_{max}}}$, since excessive output power may lead to oxygen starvation that would damage the FCS. Thus, the inequality constraint on the FCS power takes the form

$$P_{\mathrm{FC_{min}}} \le P_{\mathrm{FC}}(k) \le P_{\mathrm{FC_{max}}}. \tag{11}$$

In addition to the need to consider the magnitude of the FCS power, the rate of change of the FCS power also should be limited to ensure that the air compressor (which has a slow dynamic response) can cope with the power fluctuations. Thus, the second inequality constraint should be of the form

$$\Delta P_{\mathrm{FC,fall}} \le \frac{d}{dt} P_{\mathrm{FC}}(k) \le \Delta P_{\mathrm{FC,rise}}. \tag{12}$$

Theoretically, the upper bound of the SoC ($\mathrm{SoC_{max}}$) of the ESS is 1, but in practice this value is set somewhat smaller so as to provide a margin for safety. The lower bound of the SoC ($\mathrm{SoC_{min}}$) is required because the power-conversion efficiency of the buck-boost converter is rather poor when the ESS voltage is extremely low. Therefore, the third inequality constraint on the operation of FCHPS takes the form

$$\mathrm{SoC_{min}} \le \mathrm{SoC}(k) \le \mathrm{SoC_{max}}. \tag{13}$$

Based on Eqs. (7)–(13), the energy-management problem can be formulated as a finite-time, constrained optimal-control problem (see Problem 1).

Problem 1.

$$\min_{\{u(k)\}_{k=0}^{K-1}} J'(x(0),U) = \min_{\{u(k)\}_{k=0}^{K-1}} \left\{ \Phi(K,x(K)) + \sum_{k=0}^{K-1} \dot{m}_{\mathrm{H_2}}(u(k)) \cdot \Delta T \right\}, \tag{14}$$

subject to

$$x(k+1) = f'(x(k)),$$

$$h(x,u) = 0,$$

and

$$g_i(x,u) \ge 0 \quad (i = 1,2,\ldots,6),$$

where $x(k) \triangleq \mathrm{SoC}(k)$, $u(k) \triangleq P_{\mathrm{FC}}(k)$, U is a history (policy) of $u(k)$ and k denotes the final time. The state equation corresponds to Eq. (7), equality constraint $h(x,u) = 0$ corresponds to the constraint in Eq. (10) and the inequalities in Eqs. (11)–(13) are reformulated in the form of $g_i(x,u) \ge 0$ as follows:

$$g_1 = u(k) - P_{\mathrm{FC_{min}}} \ge 0,$$

$$g_2 = P_{\mathrm{FC_{max}}} - u(k) \ge 0,$$

$$g_3 = u(k) - u(k-1) - \Delta P_{\mathrm{FC,fall}} \Delta T \ge 0,$$

$$g_4 = \Delta P_{\mathrm{FC,rise}} \Delta T - u(k) + u(k-1) \ge 0,$$

$$g_5 = x(k) - \mathrm{SoC_{min}} \ge 0,$$

$$g_6 = \mathrm{SoC_{max}} - x(k) \ge 0.$$

2.3. Transformation of constrained optimal control

The penalty-function method [18] and Bellman's principle of optimality [29] can be used to transform Problem 1 into an infinite-time, unconstrained optimal-control problem (see Problem 2).

Problem 2.

$$\begin{aligned} \min_{\{u(k)\}_{k=0}^{\infty}} J(x(0),U) &= \min_{\{u(k)\}_{k=0}^{K-1}} \left\{ J^*(x(K)) + \sum_{k=0}^{K-1} L(k) \right\} \\ &= \min_{\{u(k)\}_{k=0}^{K-1}} \Big\{ J^*(x(K)) + J'(x(0),U) \\ &\quad + \frac{1}{2} \sum_{i=1}^{6} \sum_{k=0}^{K-1} s_i g_i^2(x(k),u(k),k) H(g_i) \Big\}, \end{aligned} \tag{15}$$

subject to $x(k+1) = f(x(k))$, where $J^*(x(K)) = \Phi(K,x(K))$ (see Eq. (9)),

$$H(g_i) = \begin{cases} 0, & \text{if } g_i \ge 0 \\ 1, & \text{if } g_i < 0 \end{cases},$$

and s_i is the penalty coefficient.

Substituting the equality constraint of Eq. (10) into the ultracapacitor model of Eq. (7) yields

$$x(k+1) = f(x(k)) \ = x(k) + \psi(x(k),u(k),P_L(k)), \tag{16}$$

where

$$\psi(x(k),u(k),P_L(k)) = -P_c\left(x(k), \frac{P_L(k) - u(k)\eta_1}{\eta_2}\right) \cdot \frac{\Delta T}{E_{\max}}. \tag{17}$$

The inequality constraints in Problem 1 are transformed into a set of penalty functions to form the adjoined cost function in Problem 2. This transforms the EMS problem into solving an infinite-time, unconstrained minimization problem, and its solution is expected to weakly converge towards an optimal solution to Problem 1 [18].

3. Synthesis of an EMS with the AOC algorithm

3.1. Necessary conditions for optimality

Considering the minimization problem of the adjoined cost function ($J(u)$) subject to the ultracapacitor model in Eq. (16), and introducing a costate variable ($\lambda(k)$) makes it possible to transform this problem into the minimization of the Hamiltonian without any constraint [29]. The Hamiltonian is defined as

$$H_k \triangleq L(k) + f(k)\lambda(k+1), \tag{18}$$

where $f(k) \triangleq f(k, x(k), u(k))$. The necessary conditions for optimality are obtained from the minimum principle as follows:

$$x(k+1) = f(k), \tag{19}$$

$$0 = \frac{\partial L(k)}{\partial u(k)} + \left(\frac{\partial f(k)}{\partial u(k)}\right)^T \lambda^*(k+1), \tag{20}$$

$$\lambda^*(k) = \frac{\partial L(k)}{\partial x(k)} + \left(\frac{\partial f(k)}{\partial x(k)}\right)^T \lambda^*(k+1), \tag{21}$$

$$\left(\left(\frac{\partial L(0)}{\partial x(0)}\right) + \left(\frac{\partial f(0)}{\partial u(0)}\right)^T \lambda^*(1)\right)^T \tilde{x}(0) = 0, \tag{22}$$

$$\left(\lambda^*(K) - \frac{\partial J^*(K, x(K))}{\partial x(K)}\right)^T \tilde{x}(K) = 0, \tag{23}$$

where '*' denotes an optimal value, $\tilde{x}(K)$ and $\tilde{x}(0)$ are arbitrary vectors corresponding to the variations, and $J^*(K, x(K))$ is the optimal terminal cost function. The AOC method assumes that initial state $x(0)$ is either fixed or free, and that final state $x(K)$ is free. Thus, $\lambda^*(0) = 0$ may satisfy the initial-point condition in Eq. (22), and the end-point condition of Eq. (23) gives $\lambda^*(K) = \partial J^*(K, x(K))/\partial x(K)$.

3.2. The AOC-EMS system

Solving the optimality conditions for an optimal EMS is challenging since the adjoined cost function, $J(u)$, is not in quadratic form and the dynamical constraint (i.e., state equation) is nonlinear. While linear optimal-control theory is unable to provide an analytical solution to this problem, the AOC method may be implemented with neural networks to obtain an optimal strategy. Fig. 7 shows a block diagram of the AOC-EMS system. The AOC algorithm consists of three blocks: the action network, $\hat{u}(k|w)$; the critic network, $\hat{\lambda}^\circ(k+1|\alpha)$; and the shadow-critic network, $\hat{\lambda}(k|\alpha)$; where w and α denote the synaptic weights of neural networks. The shadow-critic network estimates the optimal costate at the present time step, and the critic network projects this quantity to the next time step. The action network is responsible for the EMS while the critic and shadow-critic networks facilitate an adaptive critic mechanism to guide the improvement of the action network using the reinforcement-learning methodology. The critic and shadow-critic networks have identical topology and synaptic weights but different inputs and outputs.

The optimality conditions in Eqs. (20) and (21) may not hold away from an optimal trajectory. In fact, we have

$$\delta(k) = \frac{\partial L(k)}{\partial u(k)} + \left(\frac{\partial f(k)}{\partial u(k)}\right)^T \lambda(k+1), \tag{24}$$

and

$$\lambda^{\mathrm{D}}(k) = \frac{\partial L(k)}{\partial x(k)} + \left(\frac{\partial f(k)}{\partial x(k)}\right)^T \lambda(k+1) + \left(\frac{\partial u(k)}{\partial x(k)}\right)^T \delta(k). \tag{25}$$

It is evident that optimality requires $\delta(k) = 0$ and $\lambda^{\mathrm{D}}(k) = \lambda^*(k)$. Let l be the index for the version of the synaptic weights. Substituting the neural networks for corresponding terms in Eqs. (24) and (25) yields

$$\delta(k) = \left\{\frac{\partial L(k)}{\partial u(k)} + \left(\frac{\partial f(k)}{\partial u(k)}\right)^T \hat{\lambda}^\circ(k+1|\alpha_l)\right\}_{u(k)=\hat{u}(k|w_l)}, \tag{26}$$

and

$$\varepsilon(k) = \lambda^{\mathrm{D}}(k) - \hat{\lambda}(k|\alpha_l) = \frac{\partial L(k)}{\partial x(k)} + \left(\frac{\partial f(k)}{\partial x(k)}\right)^T \hat{\lambda}^\circ(k+1|\alpha_l) + \left(\frac{\partial \hat{u}(k|w_{l+1})}{\partial x(k)}\right)^T \delta(k) - \hat{\lambda}(k|\alpha_l), \tag{27}$$

where $\delta(k)$ denotes the action residual and $\varepsilon(k)$ is the critic error. These two quantities result from non-optimality or from approximation errors in neural networks.

3.3. Action-network improvement routine

The adjoined cost function in Problem 2 can be written as a recurrence relation:

$$J(k, x(k), U) = L(k) + J(k+1, x(k+1), U), \tag{28}$$

where U denotes a history of $u(k)$. Bellman's principle of optimality [29] is used to obtain

$$\min_{\{u(k)\}_{k=0}^{\infty}} J(k, x(k), U) = \min_{u(k)}\{L(k) + J^*(k+1, x(k+1))\}. \tag{29}$$

The action-network improvement routine is designed to minimize the cost function in the minimization operation. The gradient of the adjoined cost function with respect to w is

$$\frac{\partial J(k)}{\partial w} = \frac{\partial L(k)}{\partial w} + \frac{\partial J^*(k+1)}{\partial w} = \left(\frac{\partial u(k)}{\partial w}\right)^T \left\{\frac{\partial L(k)}{\partial u(k)} + \left(\frac{\partial f(k)}{\partial u(k)}\right)^T \lambda^*(k+1)\right\}, \tag{30}$$

where $\lambda^*(k) \triangleq \partial J^*(k, x(k))/\partial x(k)$ and $J(k) \triangleq J(k, x(k), U)$. Substituting the network output for the corresponding terms yields

$$\frac{\partial J(k)}{\partial w} = \left(\frac{\partial \hat{u}(k|w)}{\partial w}\right)^T \left\{\frac{\partial L(k)}{\partial u(k)} + \left(\frac{\partial f(k)}{\partial u(k)}\right)^T \hat{\lambda}^\circ(k+1|\alpha)\right\}_{u(k)=\hat{u}(k|w)}. \tag{31}$$

Using the gradient-descent method, the generalized delta rule for updating the action network is

$$w_{l+1} = w_l + \Delta w_l, \tag{32}$$

where

$$\Delta w_l = \rho_a \Delta w_{l-1} - \mu_a \left.\frac{\partial J(k)}{\partial w}\right|_{w=w_l}. \tag{33}$$

Substituting the gradient in Eq. (31) into Eqs. (32) and (33) yields the action-network updating rule:

$$\Delta w_l = \rho_a \Delta w_{l-1} - \mu_a \left\{\left(\frac{\partial \hat{u}(k|w)}{\partial w}\right)^T \delta(k)\right\}_{w=w_l}, \tag{34}$$

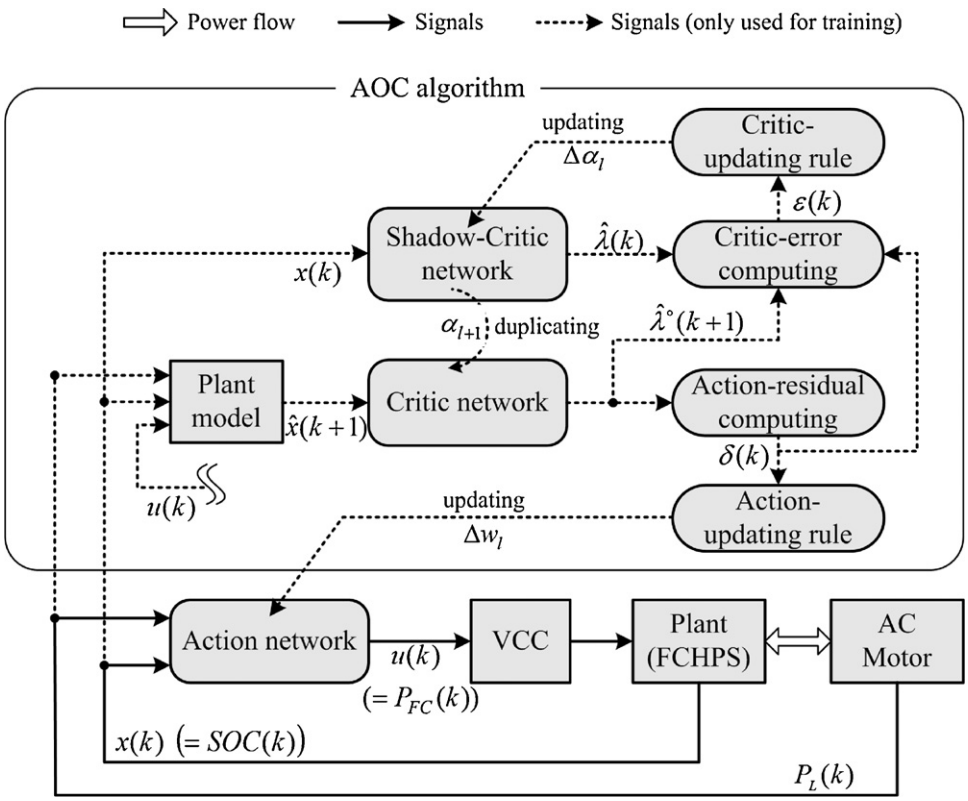

Fig. 7. The AOC-EMS system.

where ρ_a denotes the momentum coefficient and μ_a denotes the learning rate of the action network.

3.4. Critic-network learning routine

The shadow-critic network is updated using a supervised learning method by minimizing the instantaneous sum of squared critic errors at time step k:

$$E_c(k) = \frac{1}{2}\varepsilon^T(k)\varepsilon(k) = \frac{1}{2}||\lambda^{\mathrm{D}}(k) - \hat{\lambda}(k|\alpha)||_2^2. \tag{35}$$

Using the gradient-descent method yields the generalized delta rule for updating the shadow-critic network:

$$\alpha_{l+1} = \alpha_l + \Delta\alpha_l, \tag{36}$$

with

$$\Delta\alpha_l = \rho_c\Delta\alpha_{l-1} - \mu_c\left.\frac{\partial E_c(k)}{\partial\alpha}\right|_{\alpha=\alpha_l} = \rho_c\Delta\alpha_{l-1} + \mu_c\left.\frac{\partial\hat{\lambda}(k|\alpha)}{\partial\alpha}\right|_{\alpha=\alpha_l}^{T}\varepsilon(k), \tag{37}$$

where ρ_c is the momentum coefficient, μ_c is the learning rate of the shadow-critic network and gradient $\partial\hat{\lambda}(k|\alpha)/\partial\alpha$ can be obtained from the back-propagation of the shadow-critic network. The shadow-critic network receives updates stepwise (as in the instantaneous mode), while the critic network obtains a duplicate of the updated weights in turn. In calculating Eq. (27), the term $\hat{\lambda}^{\circ}(k+1|\alpha)$ is taken from the critic network for $k+1=1,\ldots,K-1$, and $\lambda^{\mathrm{D}}(K)$ at the final step is substituted by the end-point condition; that is, $\lambda^*(K) = \partial J^*(K, x(K))/\partial x(K)$.

3.5. Plant model and Jacobian quantities

As illustrated in Fig. 7, a plant model is needed to predict the step-ahead values of the plant states for computing the critic-network outputs. Using the plant model in Eq. (16) allows the Jacobian quantities appearing in Eqs. (26) and (27) to be written as follows:

$$\frac{\partial f(x(k),u(k))}{\partial u(k)} = \frac{(\eta_1\Delta T/\eta_2 E_{\max})}{\sqrt{1-(2R_{uc}C_{uc}/E_{\max})\cdot((P_L(k)-u(k)\eta_1/x(k)\eta_2)}}, \tag{38}$$

$$\frac{\partial f(x(k),u(k))}{\partial x(k)} = 1 + \frac{(P_L(k)-u(k)\eta_1/x(k)\eta_2)(\Delta T/E_{\max})}{\sqrt{1-(2R_{uc}C_{uc}/E_{\max})\cdot((P_L(k)-u(k)\eta_1)/x(k)\eta_2)}} - \left(1-\sqrt{1-\frac{2R_{uc}C_{uc}}{E_{\max}}\cdot\frac{P_L(k)-u(k)\eta_1}{x(k)\eta_2}}\right)\cdot\frac{\Delta T}{R_{uc}C_{uc}}. \tag{39}$$

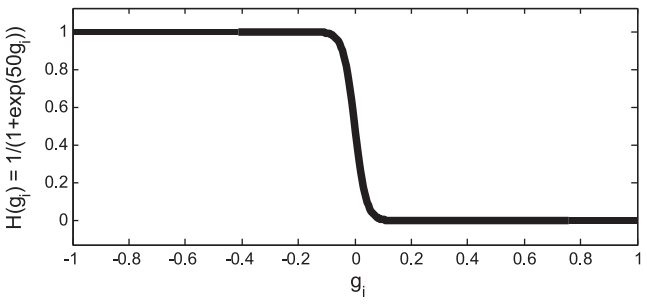

Fig. 8. A function that provides a smooth penalty transition.

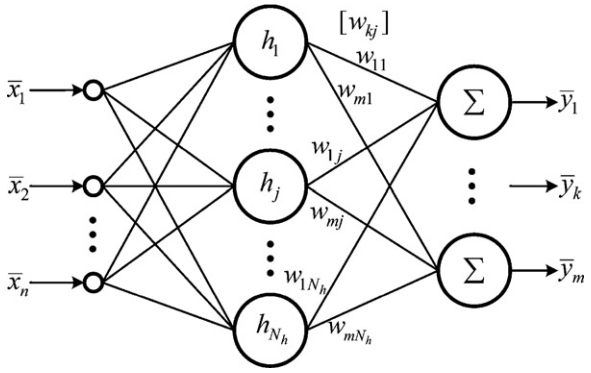

Fig. 9. Architecture of the three-layer RBF network.

The derivatives of Lagrangian $L(k)$ in adjoined cost function $J(k)$ with respect to $u(k)$ and $x(k)$ are as follows:

$$\frac{\partial L(k)}{\partial x(k)} = \frac{\partial\left(\frac{1}{2}\sum_{i=1}^{6} s_i g_i^2 H(g_i)\right)\cdot\Delta T}{\partial x(k)} = \frac{1}{2}\Delta T\cdot\sum_{i=1}^{6}\left(2s_i g_i\frac{\partial g_i}{\partial x}\cdot H(g_i) + s_i g_i^2\cdot\frac{\partial H(g_i)}{\partial g_i}\cdot\frac{\partial g_i}{\partial x}\right), \quad (40)$$

$$\frac{\partial L(k)}{\partial u(k)} = \frac{\partial \dot{m}_{H_2}}{\partial u(k)} + \frac{\partial\left(\frac{1}{2}\sum_{i=1}^{6} s_i g_i^2 H(g_i)\right)\cdot\Delta T}{\partial u(k)} = (3a_3u^2(k) + 2a_2u(k) + a_1) + \frac{1}{2}\Delta T\cdot\sum_{i=1}^{6}\left(2s_i g_i\frac{\partial g_i}{\partial u}\cdot H(g_i) + s_i g_i^2\cdot\frac{\partial H(g_i)}{\partial g_i}\cdot\frac{\partial g_i}{\partial u}\right). \quad (41)$$

$H(g_i)$ is chosen to be a sigmoid function (instead of a discontinuous step function):

$$H(g_i) = \frac{1}{1+\exp(\sigma g_i)}, \quad (42)$$

since the function must be differentiable for the generalized delta rule to be applicable, and hence must provide a smooth transition between legal and illegal regions of inequality constraints so that it is differentiable at critical point $g_i = 0$. Larger value of σ for Eq. (42) leads to a steeper slope in transition region and hence better approximation to the ideal case as defined in Problem 2. Fig. 8 shows a plot of this function for $\sigma = 50$ used in this study.

Table 2
Parameters of the electric vehicle.

Parameter	Symbol	Value
Total mass	M	2049 kg
Air drag coefficient	C_d	0.31
Frontal area	A_f	2 m^2
Air density	R_a	1.23 kg m^{-3}
Coefficient of rolling resistance	f_r	0.01
Acceleration due to gravity	g	9.8 m s^{-2}

Table 3
Parameters of the FCHPS.

Parameter	Symbol	Value
Maximum FCS power	$P_{FC_{max}}$	50 kW
Minimum FCS power	$P_{FC_{min}}$	5 kW
Maximum FCS rising power rate	$\Delta P_{FC,rise}$	5 kW s^{-1}
Maximum FCS falling power rate	$\Delta P_{FC,fall}$	−5 kW s^{-1}
Internal resistance of ultracapacitor	R_{uc}	0.035 Ω
Capacitance of ultracapacitor	C_{uc}	52 F
Maximum ultracapacitor energy	E_{max}	1625 kJ
Maximum SoC of ultracapacitor	SoC_{max}	0.95
Minimum SoC of ultracapacitor	SoC_{min}	0.25
Efficiency of boost converter	η_1	0.95
Efficiency of buck-boost converter	η_2	0.95

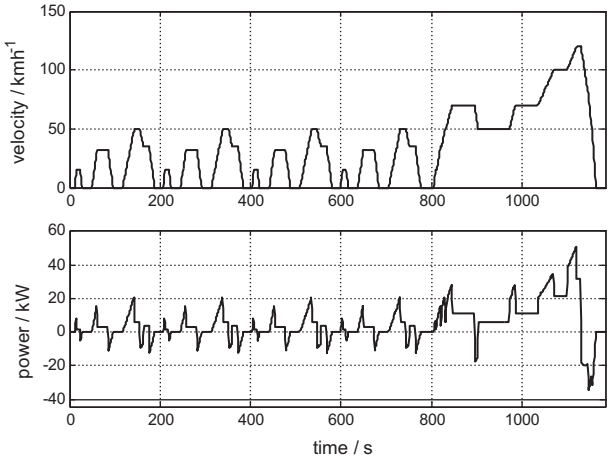

Fig. 10. Velocity time course and corresponding power demands of the NEDC.

3.6. Neural network architecture for the AOC-EMS system

In the AOC-EMS system, the associated neural networks are implemented with RBF networks [30]. Fig. 9 shows the architecture of RBF networks in which each hidden neuron h_j is modelled as a Gaussian function:

$$h_j(\bar{x}) = \exp\left(-\frac{||\bar{x}-\bar{c}_j||^2}{2b_j^2}\right), \quad (43)$$

Table 4
Fuzzy-logic rules for the EMS.

P_{ESS}	P_L				
	NB	NS	Z	PS	PB
SoC					
Z	NB	NB	NB	NS	Z
PS	NB	NS	Z	PS	PB
PB	Z	Z	Z	PS	PB

P, positive; N, negative; B, big; S, small; Z, zero.

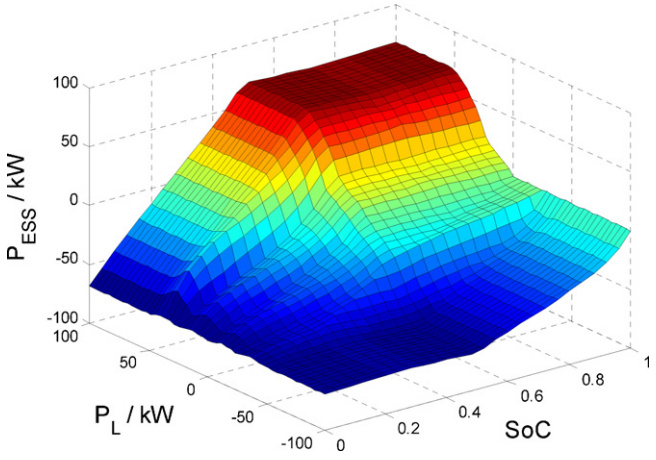

Fig. 11. Charge–discharge contour plot of the ultracapacitor bank generated by the fuzzy-logic rules.

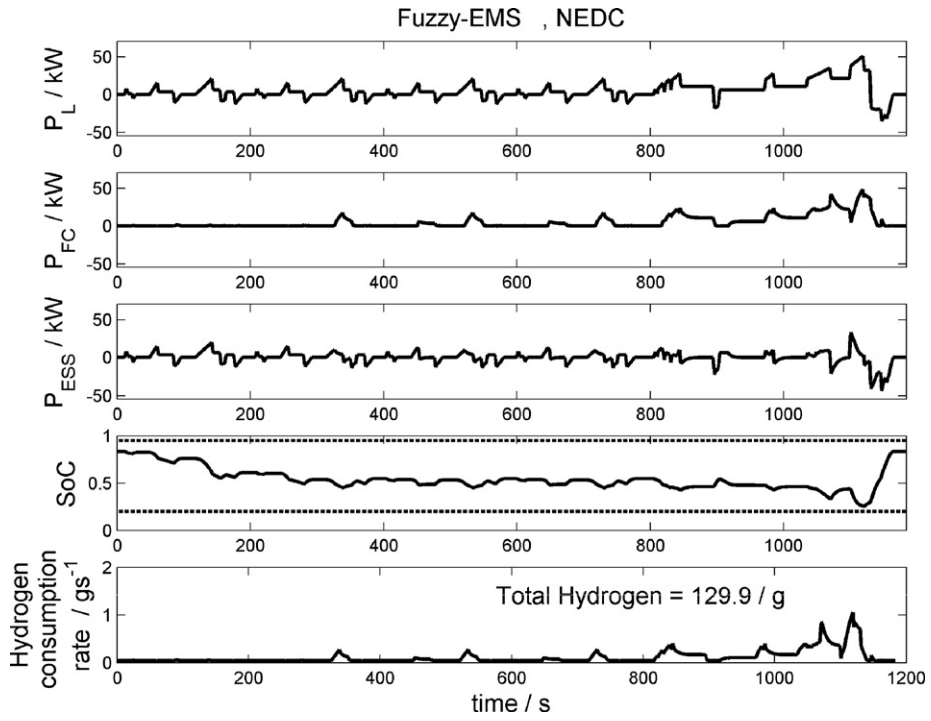

Fig. 12. Trajectories of energy management obtained by the fuzzy-EMS for the FCHPS.

where $\bar{x}$ denotes the input vector of the RBF network, $\bar{c}_j$ denotes the centre vector and b_j is the width of the Gaussian function. The functional mapping of an RBF network can be written as

$$\bar{y} = \sum_{j=1}^{N_h} \bar{w}_j h_j(\bar{x}) = \bar{w}_1 h_1(\bar{x}) + \bar{w}_2 h_2(\bar{x}) + \cdots + \bar{w}_{N_h} h_{N_h}(\bar{x}), \quad (44)$$

where $\bar{w}_j$ is a vector of synaptic weights connecting the j-th hidden neuron to output vector $\bar{y}$. The partial derivatives pertaining to the RBF network are as follows:

$$\frac{\partial \bar{y}}{\partial \bar{w}_j} = h_j(\bar{x}), \quad (45)$$

$$\frac{\partial \bar{y}}{\partial b} = \bar{w}_j \cdot h_j(\bar{x}) \cdot \frac{||\bar{x} - \bar{c}_j||^2}{b_j^3}, \quad (46)$$

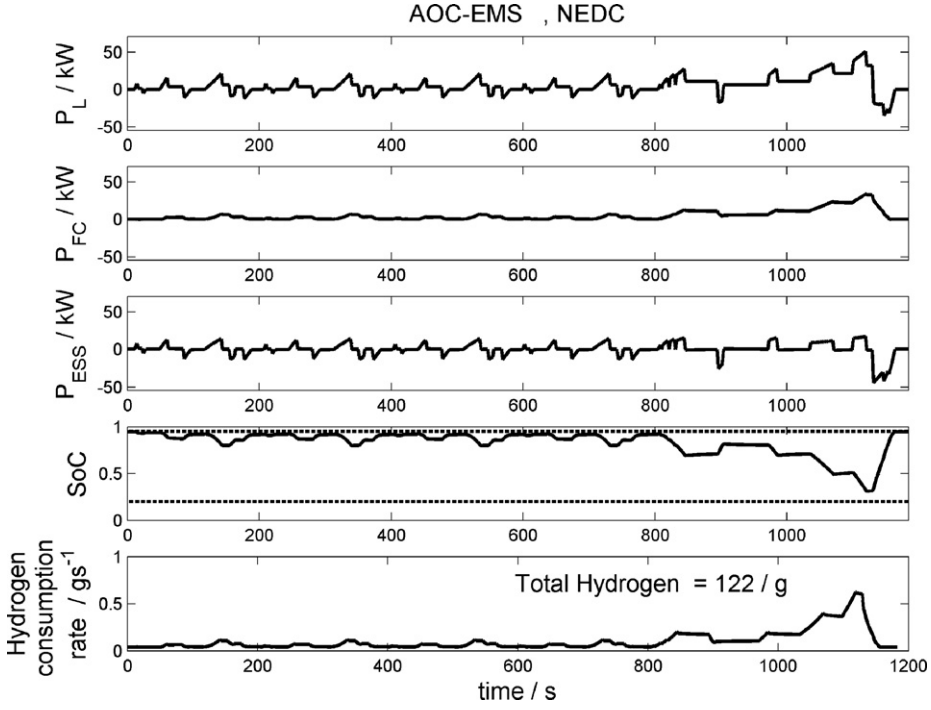

Fig. 13. Trajectories of energy management obtained by the AOC-EMS for the FCHPS.

$$\frac{\partial \bar{y}}{\partial \bar{c}_j} = \bar{w}_j \cdot h_j(\bar{x}) \cdot \frac{(\bar{x} - \bar{c}_j)^T}{b_j^2}, \tag{47}$$

$$\frac{\partial \bar{y}}{\partial \bar{x}} = -\sum_{j=1}^{N_h} \bar{w}_j \cdot h_j(\bar{x}) \cdot \frac{(\bar{x} - \bar{c}_j)^T}{b_j^2}. \tag{48}$$

These partial derivatives can be used to derive the gradient quantities such as $\partial \hat{u}(k|w)/\partial x(k)$ and $\partial \hat{\lambda}(k|\alpha)/\partial \alpha$ for the AOC-EMS system.

4. Validations

The effectiveness of the AOC-EMS system in managing the energy usage of an electric vehicle powered by an FCHPS was investigated. Table 2 lists the model parameters of the electric vehicle. The FCHPS consisted of a 50-kW FCS and a bank of ultracapacitors. With the power-train parameters listed in Table 3, the electric vehicle was capable of accelerating from 0 to 100 km in 10 s if the initial SoC of the ESS was greater than 70%. The operational limitations of the power devices, stated as [min, max], were as follows: [5, 50] kW for the fuel-cell power, $[-5, 5]\,\mathrm{kW\,s^{-1}}$ for the power-variation rate of the fuel cell and [0.2, 0.95] for the SoC of the ESS bank.

The AOC-EMS system was tested in five standard driving cycles: the New European Driving Cycle (NEDC), the UDDS (Urban Dynamometer Driving Schedule), the Highway Fuel Economy Test (HWFET), Aggressive Driving Cycle (US-06), and the FTP (Federal Test Procedure). Fig. 10 illustrates the velocity time course for the NEDC and the corresponding power demands.

For comparison, a fuzzy-logic-based EMS (fuzzy-EMS) with linguistic rules constructed by an expert was also investigated for the same driving cycles. Depending on the load power, $P_L(k)$ and SoC(k) of the ESS, the fuzzy-EMS performed a linguistic inference based on the rules listed in Table 4 to generate the correct power-split command of the ESS bank. The principal concept underlying the applied fuzzy rules was ensuring that the fuel cell satisfied the power demands while maintaining the SoC of the ESS bank within the permitted range. The charge–discharge contour plot of the ESS bank generated by the fuzzy-EMS is shown in Fig. 11.

Figs. 12 and 13 show the test results for the fuzzy-EMS and AOC-EMS, respectively. A fair comparison was ensured by requiring each test case to meet the condition that the final SoC after completing a cycle was within ±0.5% of the initial SoC. The total hydrogen consumption over the NEDC was 129.9 g for the fuzzy-EMS and 122 g for the AOC-EMS. The results show that the AOC-EMS not only minimized the hydrogen consumption but also significantly reduced both the peak power and the power-variation rate of the fuel cell.

The results of applying the designed strategy in the five driving cycles are compared with those for the electric vehicle powered by a fuel cell alone in Fig. 14. The average improvement in hydrogen consumption was 23.3% for the AOC-EMS and 18.7% for the fuzzy-EMS.

The performance of the synthesized AOC-EMS was evaluated further. Fig. 15 illustrates the fuel-cell power distribution with the AOC-EMS for three types of driving cycle. In the NEDC, which consists of frequent accelerations and decelerations, most of the fuel-cell power was used in the low-power region, meaning that the ultracapacitor bank can greatly minimize the burden on the FCS in an urban-like driving pattern. In the HWFET the vehicle is driven in a highway-like pattern, which leads to a higher sustained demand for electrical power; the results show that the power distribution shifted towards the high-efficiency region of the fuel-cell efficiency, corresponding to a higher average power. Under the US-06 test, which contains both urban-like and highway-like driving patterns, the fuel-cell power distribution was concentrated in the region characterized by extremely low hydrogen consumption and highest efficiency. These observations confirm that the AOC-EMS is indeed effective in minimizing fuel consumption under various types of driving patterns.

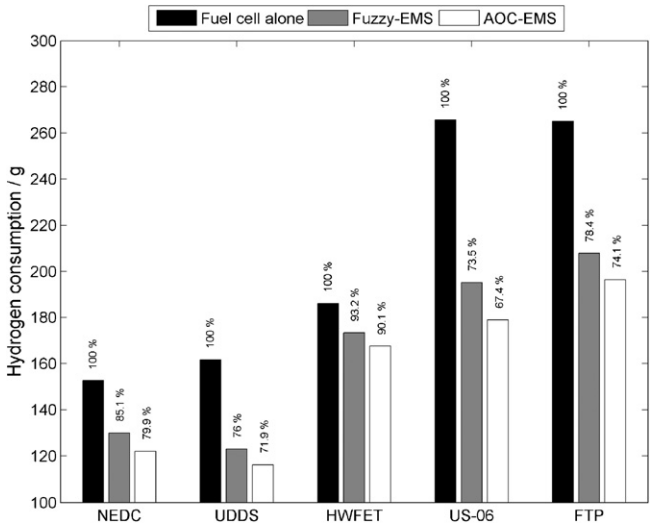

Fig. 14. Comparison of the hydrogen consumptions in five standard driving cycles.

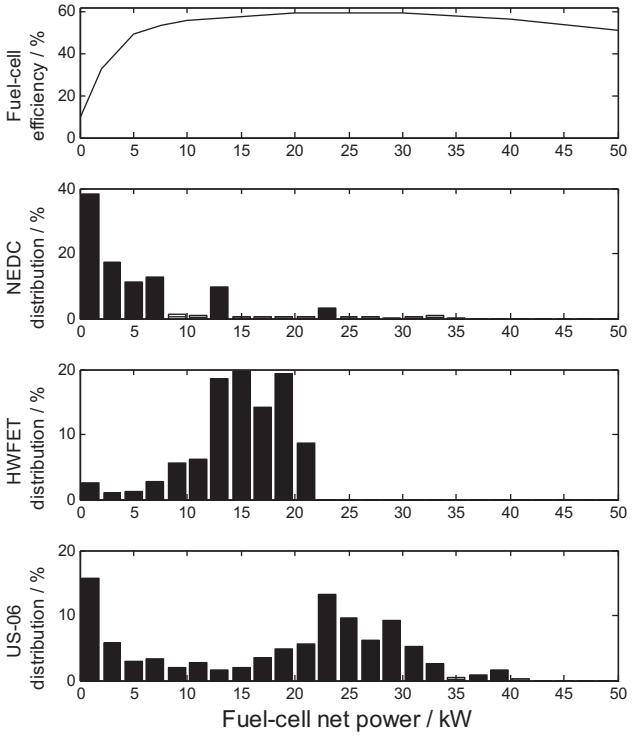

Fig. 15. Fuel-cell net-power distribution with AOC-EMS applied in the NEDC, HWFET and US-06.

5. Conclusions

This work formulated the complex requirements for the energy management of the FCHPS as a constrained optimal-control problem. The penalty-function method was used to systematically transform the (tedious) constrained problem into an unconstrained problem. The results have shown that the AOC-EMS system is able to synthesize an optimal EMS through reinforcement learning. The AOC-EMS system requires a pretraining procedure to obtain convergent weights for each neural network. In the pretraining phase,

the full pattern of a driving cycle should be presented to this system so that all possible driving situations are taken into account. To execute the EMS, convergent pretraining should produce an action network that constitutes an approximate optimal strategy of the optimal-control problem. The pretrained action network can be applied in a stochastic real-time execution without prior knowledge of future driving patterns. The AOC-EMS system can also fine tune the trained action network through online learning in order to adapt to a real-world environment.

Acknowledgement

This research was supported by National Taiwan University, Taiwan, under grant 99R30401.

References

[1] Htac, Annual Report of The Hydrogen and Fuel Cell Technical Advisory Committee (2009).
[2] L. Gao, Z. Jiang, R.A. Dougal, IEEE Transactions on Aerospace and Electronic Systems 41 (January (1)) (2005) 346–354.
[3] W. Gao, IEEE Transactions on Vehicular Technology 54 (May (3)) (2005) 846–855.
[4] A. Vahidi, A. Stefanopoulou, H. Peng, IEEE Transactions on Control Systems Technology 14 (November (6)) (2006) 1047–1057.
[5] Q. Chen, L. Gao, R.A. Dougal, S. Quan, Journal of Power Sources 191 (2009) 473–482.
[6] T. Zhu, S.R. Shaw, S.B. Leeb, IEEE Transactions on Energy Conversion 21 (March (1)) (2006) 195–201.
[7] X. Zhang, C.C. Mi, A. Masrur, D. Daniszewski, Journal of Power Sources 185 (2008) 1533–1543.
[8] Z. Jiang, Journal of Power Sources 177 (2008) 231–238.
[9] O. Erdinc, B. Vural, M. Uzunoglu, Journal of Power Sources 1094 (2009) 369–380.
[10] M. Kim, Y. Sohn, W. Lee, C. Kim, Journal of Power Sources 178 (2008) 706–710.
[11] C.Y. Li, G.P. Liu, Journal of Power Sources 192 (2009) 525–533.
[12] X. Li, L. Xu, J. Hua, Journal of Power Sources 191 (2009) 542–549.
[13] G. Paganelli, T.M. Guerra, S. Delprat, J.J. Santin, M. Delhom, E. Combes, Journal of Automobile Engineering 214 (7) (2000) 705–718.
[14] P. Rodatz, G. Paganelli, A. Sciarretta, L. Guzzella, Control Engineering Practice 13 (2005) 41–53.
[15] D. Feroldi, M. Serra, J. Riera, Journal of Power Sources 190 (2009) 387–401.
[16] S. Delprat, J. Lauber, T.M. Guerra, J. Rimaux, IEEE Transactions on Vehicular Technology 53 (3) (2004) 872–881.
[17] J. Bernard, S. Delprat, F. Buechi, T.M. Guerra, IEEE Conference (2006).
[18] A.Q. Xing, Journal of Applied Mathematics and Simulation 2 (4) (1989) 251–265.
[19] P. Werbos, in: D.A. White, D.A. Sofge (Eds.), Handbook of Intelligent Control, Van Nostraind Reinhold, New York, 1992, pp. 493–525.
[20] P. Werbos, in: J. Si, A.G. Barto, W.B. Powell, D. Wunsch (Eds.), Handbook of Learning and Approximate Dynamic Programming, IEEE Press Series on Computational Intelligence, 2004, pp. 3–44.
[21] D. Prokhorov, D. Wunsch, IEEE Transactions on Neural Networks 8 (1997) 997–1007.
[22] K.B. Wipke, M.R. Cuddy, S.D. Burch, IEEE Transactions on Vehicular Technology 48 (6) (1999) 1751–1761.
[23] J. Moreno, M.E. Ortúzar, J.W. Dixon, IEEE Transactions on Industrial Electronics 53 (April (2)) (2006) 614–623.
[24] M. Uzunoglu, M.S. Alam, IEEE Transactions on Energy Conversion 23 (March (1)) (2008) 263–272.
[25] X. Zhang, C.C. Mi, A. Masrur, D. Daniszewski, Journal of Power Sources 185 (2008) 1533–1543.
[26] J. Alvarez-Ramirez, G. Espinosa-Perez, D. Noriega-Pineda, Proceedings of the 2001 IEEE International Conference on Control Applications 5–7 (September) (2001) 190–195.
[27] L. Solero, A. Lidozzi, A. Pomilio, IEEE Transactions on Power Electronics 20 (September (5)) (2005) 1007–1016.
[28] L.K. Yi, J. Zhao, D. Ma, IEEE International Conference on Control and Automation May (2007) 1198–1201.
[29] F.L. Lewis, V.L. Syrmos, Optimal Control, second ed., John Wiley & Sons, Inc., 1995.
[30] S. Haykin, Neural Networks a Comprehensive Foundation, second ed., Prentice Hall International, Inc., 1999, pp. 156–169.