Energies **2015**, *8*, 5053-5073; doi:10.3390/en8065053

OPEN ACCESS

energies

ISSN 1996-1073

www.mdpi.com/journal/energies

Article

Dynamic Power Management for Portable Hybrid Power-Supply Systems Utilizing Approximate Dynamic Programming

Jooyoung Park [1,*], Gyo-Bum Chung [2], Jungdong Lim [1] and Dongsu Yang [3]

[1] Department of Control and Instrumentation Engineering, Korea University, Sejong 339-700, Korea; E-Mail: parkj@korea.ac.kr

[2] Department of Electronic & Electrical Engineering, Hongik University, Sejong 339-701, Korea; E-Mail: gbchung@hongik.ac.kr

[3] LG Electronics, Seoul 135-860, Korea; E-Mail: dongsu1234.yang@lge.com

* Author to whom correspondence should be addressed; E-Mail: parkj@korea.ac.kr; Tel.: +82-44-860-1444; Fax: +82-44-860-1587.

Academic Editor: William Holderbaum

Received: 17 March 2015 / Accepted: 12 May 2015 / Published: 29 May 2015

Abstract: Recently, the optimization of power flows in portable hybrid power-supply systems (HPSSs) has become an important issue with the advent of a variety of mobile systems and hybrid energy technologies. In this paper, a control strategy is considered for dynamically managing power flows in portable HPSSs employing batteries and supercapacitors. Our dynamic power management strategy utilizes the concept of approximate dynamic programming (ADP). ADP methods are important tools in the fields of stochastic control and machine learning, and the utilization of these tools for practical engineering problems is now an active and promising research field. We propose an ADP-based procedure based on optimization under constraints including the iterated Bellman inequalities, which can be solved by convex optimization carried out offline, to find the optimal power management rules for portable HPSSs. The effectiveness of the proposed procedure is tested through dynamic simulations for smartphone workload scenarios, and simulation results show that the proposed strategy can successfully cope with uncertain workload demands.

Keywords: hybrid power-supply systems; dynamic power management; approximate dynamic programming

---

1. Introduction

Recently, mobile technologies have advanced to a remarkable extent, and power consumption from mobile devices has become a serious concern in portable electronics such as smartphones, tablet PCs, and electric vehicles. The most widely used power-supply systems for portable electronic devices are electrochemical-battery-based. However, battery technology has progressed very slowly and suffers from several drawbacks such as the high current load that rapidly depletes the battery's energy and its intrinsic low power density characteristic, resulting in inefficiency when handling high load fluctuations. One of the promising methods that can compensate for some of the battery's drawbacks is based on the use of supercapacitors [1]. The supercapacitor is an emerging technology that is currently experiencing rapid development and is particularly attractive for portable systems. This is because the supercapacitor has a higher power density and faster charge/discharge capability compared to a power supply based only on batteries. Owing to the strengths of supercapacitors that can compensate for the weaknesses of batteries, portable hybrid power-supply systems (HPSSs) employing both batteries and supercapacitors are now considered to be an important alternative to conventional battery-only power supply systems [2,3]. Because the lifetime of an HPSS can be significantly extended by judiciously controlling the level of each component's power consumption, optimization of the power flows in HPSSs is a key issue that should be solved for intelligent power management of HPSSs [4]. In this paper, the problem of dynamically managing power flows in portable HPSSs employing batteries and supercapacitors is considered, and we propose an approximate dynamic programming (ADP)-based solution. In recent years, stochastic optimal control theory has been applied to various engineering problems. As is well known, a general stochastic optimal control strategy can be derived, in principle, utilizing the concept of the state value function, which is defined as the minimum total expected cost achieved by an optimal control policy from the given state. This value-function-based method is called dynamic programming (DP) and provides an important theoretical foundation for various optimal control problems. However, most real-world optimal control problems cannot be solved by DP exactly, except in very simple cases, and many studies rely on ADP, which finds an approximate solution to the stochastic optimal control problem utilizing an approximate value function, in order to obtain a good suboptimal control policy [5–9]. The main objective of this paper is to consider an ADP-based solution to dynamic power management (DPM) for portable HPSSs employing batteries and supercapacitors. More specifically, we consider a mathematical formulation of DPM for an HPSS in the form of a stochastic optimal control problem and establish an ADP-based procedure for solving the resulting controller synthesis problem. Simulation results show that this procedure works well when applied to DPM for smartphones under various load profiles.

The remainder of this paper is organized as follows. In Section 2, a preliminary background is provided regarding the mathematical formulation of the DPM problem for portable HPSSs employing batteries and supercapacitors. In Section 3, we present the framework for stochastic optimal control and describe the ADP-based controller synthesis procedure for the DPM problem. In Section 4, the effectiveness of the ADP-based procedure is illustrated by using actual workload scenarios for mobile phones. Finally, in Section 5, concluding remarks are presented.

2. Problem Formulation

This section describes the portable HPSS, and presents a mathematical model in the form of a state-space description.

2.1. System Configuration and Related Background

Recently, the optimization of power flows in HPSSs has become an important issue for the purposes of energy savings, lifetime lengthening, and so forth. For the current state-of-the-art works on the subject, the reader is referred to papers such as [10–15]. Romaus *et al.* [10] addressed the problem of controlling the power flows of a hybrid energy storage system combining batteries and double layer capacitors, and proposed the use of self-optimization methods that can consider various operating conditions and take into account several partly conflicting objective functions. As an application domain, the reference paper [10] considered an autonomous rail-bound vehicle. Chen *et al.* [11] proposed a scheduling strategy for a hybrid energy storage system for a wind farm based on predictive wind speed and the particle swarm optimization (PSO) algorithm. The strengths of the proposed scheduling strategy are that the optimization results can not only satisfy the requirements of output power quality, but also have good prediction accuracy for the wind power output. Choi *et al.* [12] presented an optimal energy management scheme based on the so-called the multiplicative-increase-additive-decrease (MIAD) policy for hybrid energy storage systems with the objectives of minimizing the magnitude/fluctuation of the battery current, and minimizing the supercapacitor energy loss. Gee *et al.* [13] analyzed and implemented a battery lifetime extension strategy for a battery/supercapacitor hybrid energy storage system, in which current filtering using a low pass filter was proposed to divert the high-frequency component of the HPSS system charge/discharge current to the supercapacitor in real time. The results of [13] showed the potential improvement in battery lifetime of HPSS in a wind power system without applying any optimization methods. Blanes *et al.* [14] considered the implementation of a hybrid energy storage system using ultracapacitor, and investigated a simple FPGA control method with a bidirectional buck-boost converter for an electrical vehicle HPSS system. In the method in [14], the battery supplies the low frequency part and the supercapacitor supplies the high-frequency part of the motor current, which reduces the battery stress and softens the battery peak current. Min *et al.* [15] presented a linear programming (LP) based method to address the problem of minimizing the installation cost while meeting the grid connection requirement, together with a quadratic programming-based approach to transmit the optimal operational signals to battery energy storage system and electric double layer capacitor modules.

This paper is concerned with an ADP-based DPM of an HPSS employing batteries and supercapacitors, which are used for portable and low-power applications. A typical example is shown in Figure 1. In this example, the HPSS consists of a single battery, two supercapacitors, two DC-DC converters, and five bidirectional switches. As is well known, a battery is currently the most widely used device that supplies energy to portable systems and has a high-energy and low-power density. However, the intrinsic low-power density of a battery often results in inefficiency, especially when handling high load fluctuations. On the other hand, supercapacitors have a low-energy and high-power density. Thus, batteries and supercapacitors can compensate for each other's weaknesses quite

well [16]. In the portable HPSS in Figure 1, the battery and supercapacitors can operate in the charging and discharging modes, respectively, or all together. DC-DC converter #1 between the battery/supercapacitors and the load of the HPSS regulates the load voltage at a constant value. DC-DC converter #2 between the battery and the supercapacitors controls the charging current, which is calculated by the DPM unit. Because energy loss may occur during the process of DC-DC conversion, the converter topologies need be carefully selected [17]. Switches connect the components of the portable HPSS following the operation strategy provided by DPM. The DPM monitors the system at each sampling time step and determines the optimal decision to lengthen the lifetime of the HPSS, thus improving the energy conversion efficiency.

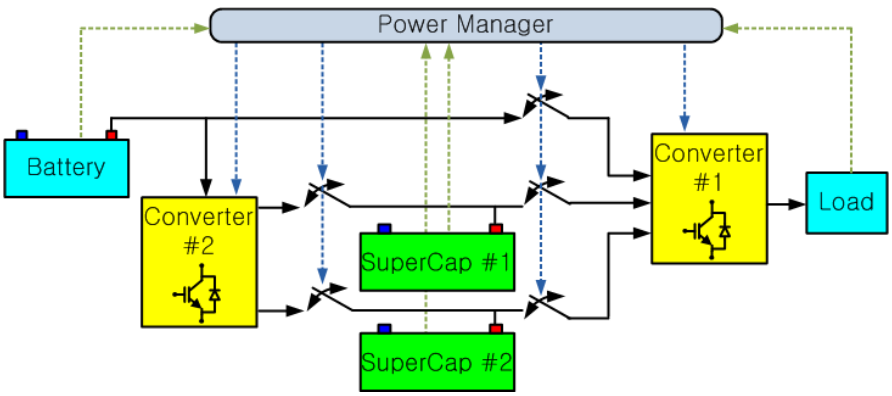

Figure 1. Structure of a typical portable hybrid power supply system.

2.2. State Equation and Performance Index

In this section, we describe the DPM of a portable HPSS employing batteries and capacitors within the framework of a stochastic optimal control problem. Throughout this paper, we consider an HPSS with the structure shown in Figure 1, which has a single battery and two supercapacitors. Generalization of the model and the corresponding solution procedure for a case with any multiple number of batteries and supercapacitors is straightforward. The steps for deriving the state equation and the performance index (PI) for the system are based on the results of [18] with some modifications. The details of these modifications will be specified later in this section.

Generally, dynamical systems in discrete-time stochastic optimal control problems are described by the state-space model in the form of:

$$x(t+1) = f(x(t), u(t), w(t)), \;\; t = 0,1,..., \tag{1}$$

where $x(t) \in X$ is the state vector, $u(t) \in U$ is the control input vector; and $w(t)$ is the external disturbance, all at time t. In the DPM problem presented in this paper, $w(t)$ is the load demand during a single time interval $(t, t + 1)$. The state vector that we consider for the DPM problem contains the amounts of charge of the supercapacitors and battery as well as the change in the battery's charge, *i.e.*,

$$x(t) = \begin{bmatrix} r_{cap,1}(t) \\ r_{cap,2}(t) \\ r_{bat}(t) \\ \Delta r_{bat}(t) \end{bmatrix}, \tag{2}$$

where $r_{cap,i}(t)$ is the charge state of the ith supercapacitor at time t, $r_{bat}(t)$ is the charge state of the battery at time t, and $\Delta r_{bat}(t) = r_{bat}(t-1) - r_{bat}(t)$ is the difference in the charge state of the battery. Note that in this paper, we assume that the state-of-charge (SOC) values of the supercapacitors and the battery are all available. In practical implementations, these SOC values can be estimated by e.g., extended Kalman filters [19]. The control input vector we consider for the DPM problem contains information regarding the proportion of the charge that each component (*i.e.*, the ith supercapacitor or battery) supplies to meet the workload and the level of the battery charge transferred to each supercapacitor. As in [18], an upper bound R_{th} is placed on the amount of charge that can be transferred from a battery to each supercapacitor within a time interval so that the battery charge may not be exhausted during the process of charging the supercapacitors. More specifically, the control input vector consists of the following five entries, all in the range of [0, 1]:

$$u(t) = \begin{bmatrix} a_{1x}(t) \\ a_{1y}(t) \\ a_{2x}(t) \\ a_{2y}(t) \\ a_{bat}(t) \end{bmatrix} \tag{3}$$

Here, $a_{ix}(t)$, $i = 1, 2$, is the normalized level of the ith supercapacitor being charged by the battery; thus, the amount of charge transferred from the battery to the ith supercapacitor during the time interval $(t, t+1)$ is $R_{th} \cdot a_{ix}(t)$. The other entries, $a_{1y}(t)$, $a_{2y}(t)$, and $a_{bat}(t)$, are the proportions of the charges that the two supercapacitors and battery supply to meet the workload $w(t)$ in the time interval $(t, t+1)$. Because these entries are all proportions, they satisfy:

$$a_{1y}(t) + a_{2y}(t) + a_{bat}(t) = 1 \text{ for } t = 0, 1, \ldots \tag{4}$$

With the above definitions of the state and control inputs, the state transition from time step t to $(t+1)$ can be expressed as follows [18]:

$$r_{cap,1}(t+1) = r_{cap,1}(t) - a_{1y}(t) w(t) + a_{1x}(t) R_{th} \tag{5}$$

$$r_{cap,2}(t+1) = r_{cap,2}(t) - a_{2y}(t) w(t) + a_{2x}(t) R_{th}, \tag{6}$$

$$r_{bat}(t+1) = r_{bat}(t) - a_{bat}(t) w(t) - a_{1x}(t) R_{th} - a_{2x}(t) R_{th}, \tag{7}$$

$$\begin{aligned} \Delta r_{bat}(t+1) &= r_{bat}(t) - r_{bat}(t+1) \\ &= a_{bat}(t) w(t) + a_{1x}(t) R_{th} + a_{2x}(t) R_{th}. \end{aligned} \tag{8}$$

Note that with the above definitions of the state vector and control input vector, these state transitions can be compactly written in the following state equation form:

$$\begin{aligned} x(t+1) &= Ax(t) + B(t)u(t) \\ &= Ax(t) + \left(B_1 + B_w w(t)\right)u(t), \end{aligned} \tag{9}$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{10}$$

$$B_1 = R_{th} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{11}$$

and

$$B_w = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{12}$$

In addition, note that $u(t)$ should satisfy the following constraints from the definitions of the control inputs:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \end{bmatrix} u(t) = 1, \tag{13}$$

$$0_{5\times1} \le u(t) \le 1_{5\times1} \tag{14}$$

In real portable HPSS applications, exhausting the charge of any component during operation is not desirable. Considering this practical constraint, one can additionally place constraints on the operating ranges of the battery and supercapacitors as follows:

$$R_{\min}^{cap} \le r_{cap,1}(t) \le R_{\max}^{cap} \tag{15}$$

$$R_{\min}^{cap} \le r_{cap,2}(t) \le R_{\max}^{cap} \tag{16}$$

$$R_{\min}^{bat} \le r_{bat}(t) \tag{17}$$

As mentioned before, the state equation is based on the HPSS model of Mirhoseimi and Koushanfar in [18]. The main differences between these two models are as follows:

- In [18], it is assumed that the charge values of the batteries and supercapacitors take only discrete values. In this study, we omit this assumption; thus, $r_{cap,i}(t)$ and $r_{bat}(t)$ are all real-valued.
- In the model in [18], supercapacitors are constrained to not be simultaneously charged by the battery and discharged by the load. In this paper, we omit this constraint.
- In the model in [18], a decision for assigning the source to the workload is carried out such that only one electronic energy supply source can transfer the required charge to the load. Hence, the control inputs in [18] are all binary numbers, and only one member of $a_{iy}(t)$ and $a_{bat}(t)$ is one. In this study, we omit this constraint. As a result, $a_{iy}(t)$ and $a_{bat}(t)$ are nonnegative real numbers satisfying $\sum_i a_{iy}(t) + a_{bat}(t) = 1$.
- In the model in [18], it is assumed that at most one supercapacitor can be charged by the battery. This assumption is omitted in this paper.

In order to solve the DPM problem within the framework of stochastic optimal control, we have to choose a PI that can ensure improved energy savings after optimization. A widely used choice for the PI in a stochastic optimal control problem is the expected sum of the discounted stage costs, *i.e.*,

$$J = E\left[\sum_{t=0}^{T} \gamma^t \ell\left(x(t), u(t)\right)\right] \tag{18}$$

where $\gamma \in (0, 1)$ is the discount factor; $l(\cdot,\cdot)$ is the stage cost function; and T is the final time whose value is ∞ in infinite-horizon problems. By minimizing this PI over all admissible state feedback control policies ϕ_t: $X \rightarrow U$, the optimal control problem can be solved. For the performance index in the optimization of power flows for HPSSs, one may resort to some ad-hoc stage cost functions (e.g., [18]). However in this paper, it turns out that a traditional objective function which is more clear and widely used in the field of linear quadratic optimal control works well. More precisely, we derive an ADP-based solution to the dynamic power management of HPSSs with the purpose of minimizing the battery charge consumption rate and maintaining the charge level of supercapacitors while keeping the control efforts reasonably low. For the purpose, the stage cost function $l(\cdot,\cdot)$ of the DPM task is chosen as follows:

$$\begin{aligned} l\left(x(t), u(t)\right) = {} & w_1 \Delta r_{bat}(t) \\ & + w_2 \sum_{i=1}^{2} \left\{ \left(r_{cap,i}(t) - 0.5 R_{\max}^{cap} \right)^2 \right\} \\ & + w_3 \left\{ \sum_{i=1}^{2} a_{ix}^2(t) \right\} + w_4 \left\{ \left(\sum_{i=1}^{2} a_{iy}^2(t) \right) + \beta_{bat} \left(a_{bat}^2(t) \right) \right\}, \end{aligned} \tag{19}$$

where the weight values (w_1, w_2, w_3, w_4, and β_{bat}) are obtained empirically via a tuning process. In the DPM in [18], the Q-learning method [20] was used to derive an approximation for the optimal power management strategy, and the stage function was defined as

$$R\left(x(t), u(t)\right) = -e^{\Delta r_{bat}(t)} - e^{\left(\sum_i r_{cap,i}(t)/R_{\max}^{cap}\right)} - \lambda e^{E\left[\Delta r_{bat}(t)\right]}. \tag{20}$$

Here, we adopt the main idea of this stage reward function with the following significant modifications. First, representing the stage costs in a quadratic form is important for solving the DPM problem in the context of constrained linear quadratic stochastic optimal control. Hence, our stage cost function utilizes second order polynomial terms instead of exponential functions. Second, in order to minimize the battery charge consumption rate and maintain the charge level of supercapacitors with the control efforts kept reasonably low, we employ the cost terms involving $\Delta r_{bat}(t)$, $(r_{cap,i}(t) - 0.5R_{\max}^{cap})^2$, $i = 1, 2$, and the control-input-related quadratic terms, $a_{ix}^2(t)$, $a_{iy}^2(t)$, $i = 1, 2$, and $a_{bat}^2(t)$. Finally, note that our stage cost Equation (19) deals with the control effort directly with weighted summation of the quadratic input terms, whereas in the objective Equation (20) of the reference paper [18], consideration of the control efforts can be treated indirectly via the expectation operation.

3. Approximate Dynamic Programming Approach to Dynamic Power Management

3.1. Preliminaries

General stochastic optimal control problems can be analyzed, in principle, utilizing the concept of a state value function. This value-function-based method is called DP, and a variety of topics on stochastic optimal control and DP are well-addressed by [5–8]. A large class of stochastic optimal control problems deal with the dynamics of the form in Equation (1) and are concerned with finding a state-feedback control policy:

$$u(t) = \phi_t\left(x(t)\right),\ t = 0,1,\ldots \tag{21}$$

that can optimize a given PI function. As mentioned before, a widely used choice for the PI is the expected sum of the discounted stage cost, *i.e.*, $J = E\left[\sum_{t=0}^{\infty}\gamma^t \ell\left(x(t),u(t)\right)\right]$. By minimizing this PI over all admissible control policies ϕ_t: $X \to U$, the optimal value of J can be found. This minimal PI value is denoted J^*, and the optimal state-feedback function achieving the minimal value is denoted by ϕ_t^*. The state value function is defined as the minimum total expected cost achieved by an optimal control policy from the given initial state $x(0) = z$, *i.e.*,

$$V^*(z) = \inf_{\phi_t} E\left(\sum_{t=0}^{\infty}\gamma^t \ell\left(x(t),\phi_t\left(x(t)\right)\right) \mid x(0) = z\right) \tag{22}$$

Note that when the initial condition $x(0)$ is a fixed vector x_0, the optimal PI value can be expressed as $J^* = V^*(x_0)$. According to stochastic optimal control theory [6,7], the state value function in Equation (23) is the fixed point of the Bellman equation:

$$V^*(z) = \inf_v\left(l(z,v) + \gamma E\left[V^*\left(f(z,v,w)\right)\right]\right). \tag{23}$$

In its operator form, the Bellman equation can be written as:

$$V^* = TV^* \tag{24}$$

where T is the Bellman operator (see e.g., [5]), whose domain and codomain are both function spaces mapping X onto $R \cup \{\infty\}$, defined as:

$$(Tg)(z) = \inf_v(\ell(z,v) + \gamma E\left[g(f(z,v,w))\right] \tag{25}$$

As is well known, the Bellman equation or its operator equation version cannot be solved exactly, except in simple special cases [5,6], and an alternative strategy when finding the exact state value function is the use of ADP relying on an approximate state value function $\hat{V}: X \to R$. In the DPM problem discussed in this paper, we are concerned with an ADP solution utilizing convex quadratic functions. This class of quadratic functions is a good choice for the approximate state value function [5,8] because these functions are convenient and powerful for handling various optimization steps appearing in the ADP procedure.

3.2. ADP-Based Solution Procedure

From the above preliminary steps, the DPM of the portable HPSS can now be represented as the following dynamic stochastic control problem:

$$\begin{aligned} \min_{u(\cdot)} \quad & E\left[\sum_{t=0}^{\infty}\gamma^t l\left(x(t),u(t)\right)\right] \\ \text{subject to} \quad & x(t+1)=Ax(t)+\left(B_1+B_w w(t)\right)u(t), \\ & \text{constraints } (13),(14),(15),(16),(17), \\ & \text{and } x(0)=x_0, \end{aligned} \tag{26}$$

where $l(\cdot,\cdot)$ is the quadratic stage cost function specified in Equation (19). To solve the above DPM problem for the portable HPSS, we propose an ADP-based solution procedure utilizing the approximate value function (AVF) policy approach of O'Donoghue, Wang, and Boyd [9]. In this AVF policy approach [9], the convex quadratic function:

$$\hat{V}_i(x)=x^T P_i x+2p_i^T+q_i,\ i=0,\ldots,M \tag{27}$$

is used for approximating the state value functions, and by letting the parameters *(i.e.*, P_i, p_i, and q_i) satisfy the iterated Bellman inequalities:

$$\hat{V}_i \le T\hat{V}_{i+1},\ i=0,\ldots,M \tag{28}$$

with $\hat{V}_M=\hat{V}_{M+1}$, it is ensured that $\hat{V}_0$ is a lower bound of the optimal state value function V^* [5,9]. Further, by optimizing this lower bound $\hat{V}_0$ via convex optimization, the ADP approach finds a suboptimal approximate state value function $\hat{V}_0,\hat{V}_1,\cdots,\hat{V}_M$ and the corresponding ADP policies $\hat{\phi}_0,\hat{\phi}_1,\cdots,\hat{\phi}_M$. In the following, we derive an ADP-based solution to the dynamic management problem for the portable HPSS.

First, note that our stage cost function at time ***t*** can be expressed using the state $x(t)$ and input $u(t)$ as follows:

$$\begin{aligned} l\left(x(t),u(t)\right)=\ & w_1\Delta r_{bat}(t)+w_2\sum_{i=1}^{2}\left(r_{cap,i}(t)-0.5R_{\max}^{cap}\right)^2 \\ & +w_3\left(a_{1x}^2(t)+a_{2x}^2(t)\right) \\ & +w_4\left(a_{1y}^2(t)+a_{2y}^2(t)+\beta_{bat}a_{bat}^2(t)\right). \end{aligned} \tag{29}$$

Then, the general stage cost term can be written in the following matrix–vector form:

$$\begin{aligned} l(x,u)&=\begin{bmatrix}u\\x\\1\end{bmatrix}^T L\begin{bmatrix}u\\x\\1\end{bmatrix} \\ &=\begin{bmatrix}u\\x\\1\end{bmatrix}^T\begin{bmatrix}L_{11} & L_{12} & L_{13}\\ L_{12}^T & L_{22} & L_{23}\\ L_{13}^T & L_{23}^T & L_{33}\end{bmatrix}\begin{bmatrix}u\\x\\1\end{bmatrix}, \end{aligned} \tag{30}$$

where

$$L_{11} = \begin{bmatrix} w_3 & 0 & 0 & 0 & 0 \\ 0 & w_4 & 0 & 0 & 0 \\ 0 & 0 & w_3 & 0 & 0 \\ 0 & 0 & 0 & w_4 & 0 \\ 0 & 0 & 0 & 0 & \beta_{bat} w_4 \end{bmatrix}, \tag{31}$$

$$L_{12} = 0_{5\times 4}, \tag{32}$$

$$L_{13} = 0_{5\times 1}, \tag{33}$$

$$L_{22} = \begin{bmatrix} w_2 & 0 & 0 & 0 \\ 0 & w_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{34}$$

$$L_{23} = \begin{bmatrix} -0.5 w_2 R_{\max}^{cap} \\ -0.5 w_2 R_{\max}^{cap} \\ 0 \\ 0.5 w_1 \end{bmatrix}, \tag{35}$$

$$\text{and } L_{33} = 0.5 w_2 \left(R_{\max}^{cap} \right)^2. \tag{36}$$

Next, we define the derived matrix variable G_i, $i = 1, \ldots, M$ to satisfy the following:

$$\begin{bmatrix} u \\ x \\ 1 \end{bmatrix}^T G_i \begin{bmatrix} u \\ x \\ 1 \end{bmatrix} = E\left[\hat{V}_i \left(Ax + \left(B_1 + B_w w \right) u \right) \right]. \tag{37}$$

Note that the right hand side of Equation (37) is the expectation of the convex quadratic function $\hat{V}_i$ evaluated at the next state. Since the expectation can be expressed as a quadratic function with respect to u and x, we use the derived matrix variable G_i [8] for notational convenience. Here, the expectation of the right-hand side of Equation (37) can be further expanded as follows:

$$\begin{aligned} & E\left[\begin{bmatrix} Ax + \left(B_1 + B_w w \right) u \\ 1 \end{bmatrix}^T \begin{bmatrix} P_i & p_i \\ p_i^T & q_i \end{bmatrix} \begin{bmatrix} Ax + \left(B_1 + B_w w \right) u \\ 1 \end{bmatrix} \right] \\ & = \begin{bmatrix} u \\ x \\ 1 \end{bmatrix} E\left[\begin{bmatrix} B_1 + B_w w & A & 0 \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} P_i & p_i \\ p_i^T & q_i \end{bmatrix} \begin{bmatrix} B_1 + B_w w & A & 0 \\ 0 & 0 & 1 \end{bmatrix} \right] \begin{bmatrix} u \\ x \\ 1 \end{bmatrix}. \end{aligned} \tag{38}$$

Furthermore, by further evaluating the right-hand side of Equation (38), one can obtain

$$G_i = \begin{bmatrix} G_{i,11} & G_{i,12} & G_{i,13} \\ \left(G_{i,12} \right)^T & G_{i,22} & G_{i,23} \\ \left(G_{i,13} \right)^T & \left(G_{i,23} \right)^T & G_{i,33} \end{bmatrix}, \tag{39}$$

where

$$G_{i,11} = B_1^T P_i B_1 + B_1^T P_i B_w \overline{w} + B_w^T P_i B_1 \overline{w} + B_w^T P_i B_w \overline{w^2}, \tag{40}$$

$$G_{i,12} = B_1^T P_i A + B_w^T P_i A \overline{w} \tag{41}$$

$$G_{i,13} = B_1^T p_i + B_w^T p_i \overline{w} \tag{42}$$

$$G_{i,22} = A^T P_i A \tag{43}$$

$$G_{i,23} = A^T p_i \tag{44}$$

$$G_{i,33} = q_i \tag{45}$$

In Equations (40)–(42), $\overline{w}$ and $\overline{w^2}$ are the first and second moments of the external workload demands, respectively.

With the convex quadratic approximation $\hat{V}_i(x) = x^T P_i x + 2p_i^T + q_i,\ i = 0,...,M$, the Bellman inequalities $\hat{V}_{i-1} \le T\hat{V}_i,\ i = 1,...,M$, for the DPM problem can be written as

$$\begin{bmatrix} u \\ x \\ 1 \end{bmatrix}^T \left(L + \gamma G_i - S_{i-1}\right) \begin{bmatrix} u \\ x \\ 1 \end{bmatrix} \ge 0,\ i = 1,...,M, \tag{46}$$

where S_{i-1} is the matrix variable defined by

$$S_{i-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & P_{i-1} & p_{i-1} \\ 0 & p_{i-1}^T & q_{i-1} \end{bmatrix}. \tag{47}$$

Note that the constraints in Equations (13)–(17) are all given in the form of an affine equality or affine inequalities with respect to the state and control inputs. Thus, these constraints can be compactly expressed via the affine form $Du(t) + Fx(t) + H$. More specifically, the equality constraint in Equation (13) can be written as

$$D^{(1)}u(t) + F^{(1)}x(t) + H^{(1)} = 0, \tag{48}$$

where $D^{(1)} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \end{bmatrix}$, $F^{(1)} = 0_{1\times 4}$, and $H^{(1)} = -1$.

Similarly, the remaining inequality constraints can be represented by

$$D^{(k)}u(t) + F^{(k)}x(t) + H^{(k)} \le 0,\ k = 2,3,4,5, \tag{49}$$

where

$$D^{(2)} = -I_{5\times 5}, \tag{50}$$

$$F^{(2)} = 0_{5\times 4}, \tag{51}$$

$$H^{(2)} = 0_{5\times 1}, \tag{52}$$

$$D^{(3)} = I_{5\times 5}, \tag{53}$$

$$F^{(3)} = 0_{5\times 4}, \tag{54}$$

$$H^{(3)} = -1_{5\times 1}, \tag{55}$$

$$D^{(4)} = 0_{3\times 5}, \tag{56}$$

$$F^{(4)} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \tag{57}$$

$$H^{(4)} = \begin{bmatrix} R_{\min}^{cap} \\ R_{\min}^{cap} \\ R_{\min}^{bat} \end{bmatrix}, \tag{58}$$

$$D^{(5)} = 0_{2\times 5}, \tag{59}$$

$$F^{(5)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \tag{60}$$

$$\text{and } H^{(5)} = \begin{bmatrix} -R_{\max}^{cap} \\ -R_{\max}^{cap} \end{bmatrix}. \tag{61}$$

Note that these constraints can be further modified in accordance with the requirements considered in the design stage. One of the key steps in this solution procedure is to find $\hat{V}_t$ satisfying the Bellman inequalities in Equation (28) under the constraints in Equations (46) and (48). For this, we resort to the S-procedure [21]. On the basis of the S-procedure, we can ensure that the Bellman inequalities be satisfied under the constraints in Equations (46) and (48) by requiring that the following linear matrix inequalities (LMIs) hold true:

$$L + \gamma G_i - S_{i-1} + \sum_k \xi_i^{(k)} \Xi^{(k)} \geq 0, \tag{62}$$

$$diag\left(\xi_i^{(k)}\right) \geq 0, \ \forall k, \ i = 1, \ldots, M, \tag{63}$$

where $\xi_i^{(k)}$ represents the S-procedure multipliers [21], and

$$\Xi^{(k)} = \begin{bmatrix} 0 & 0 & D^{(k)}/2 \\ 0 & 0 & F^{(k)}/2 \\ \left(D^{(k)}/2\right)^T & \left(F^{(k)}/2\right)^T & H^{(k)} \end{bmatrix}. \tag{64}$$

By combining the above steps, one can obtain a DPM design procedure. The following show an illustrative design procedure:

Controller synthesis procedure:

Preliminary steps:

1. Choose the parameters of the problem: γ, λ, M, R_{th}, $R_{\min}^{cap}$, $R_{\max}^{cap}$, and $R_{\min}^{bat}$.
2. Estimate the 1st and 2nd moments of the external load demands, $\overline{w}$ and $\overline{w^2}$, from the training data.

Main steps:

1. Initialize the decision-making time $t = 0$, and choose $x(0)= x_0$.
2. Compute the stage cost matrix L in Equation (30) and the constant matrix $\Xi^{(k)}$ in Equation (64).
3. Define the LMI variables:
   (1) Define the basic LMI variables: P_i, p_i, and q_i in Equation (27).
   (2) Define the derived LMI variables: G_i in Equation (39) and S_{i-1} in Equation (47).
   (3) Define the S-procedure multipliers: $\xi_i^{(k)}$ in Equation (63).
4. Find the approximate state value functions, $\hat{V}_i$, by solving the following convex optimization problem:

$$\begin{aligned} &\min \quad \hat{V}_0(x_0) = x_0^T P_0 x_0 + 2p_0^T x_0 + q_0 \\ &\text{subject to} \\ &\qquad L + \gamma G_i - S_{i-1} + \sum_{k=1}^{3} \xi_i^{(k)} \Xi^{(k)} \geq 0,\ i = 1,\ldots,M, \\ &\qquad S_{M-1} = S_M, \\ &\qquad P_i \geq 0,\ i = 0,\ldots,M, \\ &\qquad diag\left(\xi_i^{(k)}\right) \geq 0,\ i = 1,\ldots,M,\ k = 2,3. \end{aligned} \tag{65}$$

5. Obtain the ADP controllers on the basis of

$$\Phi_t(x) = \underset{u}{\operatorname{argmin}}(\ell(x,u) + \gamma E V_{t+1}(f(x,u,w))) \text{ for } t = 0,\ldots,M \tag{66}$$

$$\Phi_t(x) = \underset{u}{\operatorname{argmin}}(\ell(x,u) + \gamma E V_{T+1}(f(x,u,w))) \text{ for } t > M \tag{67}$$

4. Simulation Results and Trajectories

In this section, we first describe a portable HPSS example to illustrate the proposed ADP-based DPM procedure. Afterwards, we conduct simulations and provide discussions for simulation results.

4.1. An Illustrative Example

We conducted simulations for a portable HPSS example [18] to illustrate the proposed ADP-based DPM procedure. For the example, we considered the workload scenarios of Figure 2, which are essentially based on the iPhone battery current measurements in [18,22]. The tasks considered in the scenarios include airplane mode (40 mA), default mode (100 mA), working internet with 3G network (310 mA), working game application "Tap-Tap" (220 mA), working 3D game application "GTI Racing" (400 mA), working application "YouTube" with WiFi (260 mA), working application "YouTube" (350 mA), and voice phone (170 mA) [18].

The portable HPSS in our example consists of two 200 [F] supercapacitors with $R_{\max}^{cap}$ = 400 [C] and a single 4 [V] battery with $r_{bat}(0)$ = 1000 [C]. The initial state conditions for the HPSS were assumed as follows:

$$\begin{aligned} x(0) &= \left[r_{cap,1}(0), r_{cap,2}(0), r_{bat}(0), \Delta r_{bat}(0)\right] \\ &= \left[400, 400, 1000, 0\right]^T. \end{aligned} \tag{68}$$

Following the data treatment of [18], we rescaled the source and demand charges such that at each instant t, the load charge demand $w_{load}(t)$ was transformed into $w(t) = w_{load}(t)/10$. We performed simulations based on the ten scenarios of Figure 2, and evaluated the resultant controllers. In the evaluation stage, we used a 10-fold cross-validation method for the iPhone workload data of Figure 2. For the simulation results, we considered all possible 10 different partitions of the training and test subsets. For the parameters of the problem, we chose the following:

$$\gamma = 0.99,\ R_{th} = 1,\ R_{\min}^{cap} = 0,\ R_{\max}^{cap} = 40, R_{\min}^{bat} = 0,\ M = 150,$$
$$w_1 = 5, w_2 = 0.2,\ w_3 = 0.5,\ w_4 = 5,\ \beta_{bat} = 5,$$

where the weight values, w_1, w_2, w_3, w_4, and β_{bat} were obtained empirically via a tuning process based on a subset of the training data. Simulation results show that the above set of weight values yielded reasonably good results. To develop a more disciplined design guideline for the weight values is an important topic that can be covered in follow-up works.

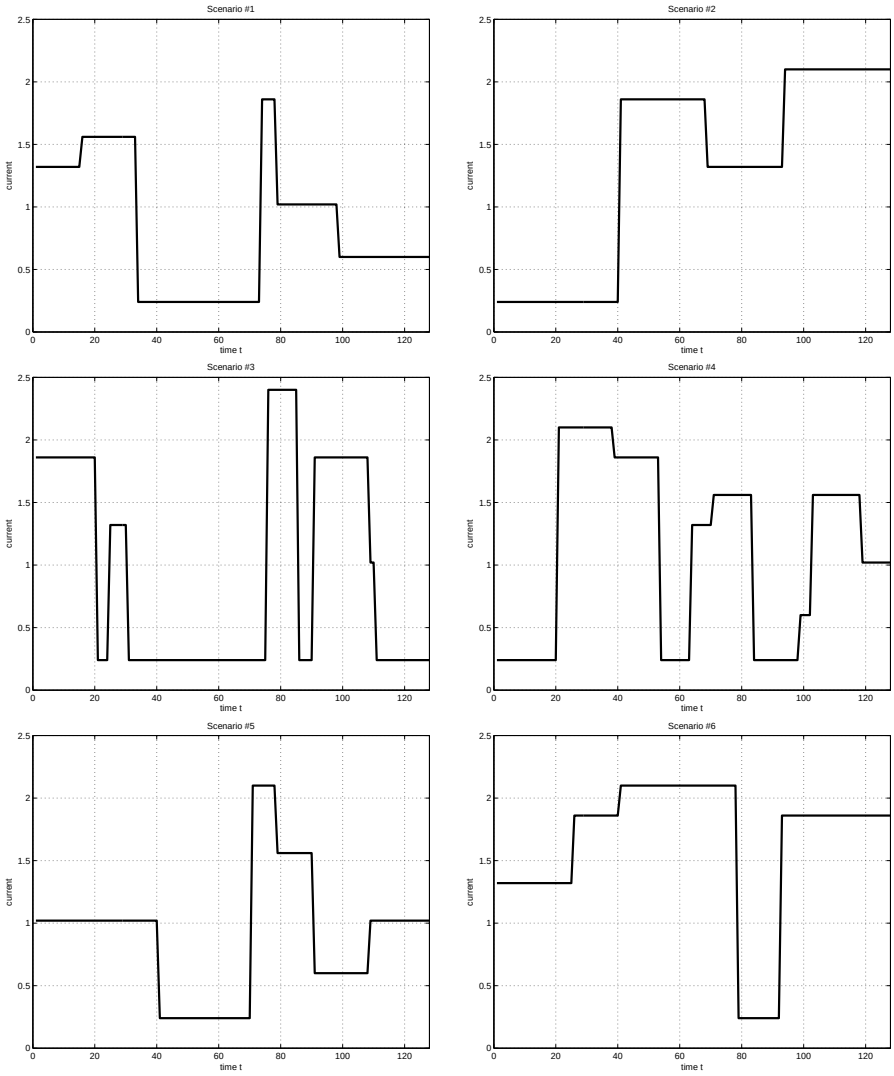

Figure 2. *Cont.*

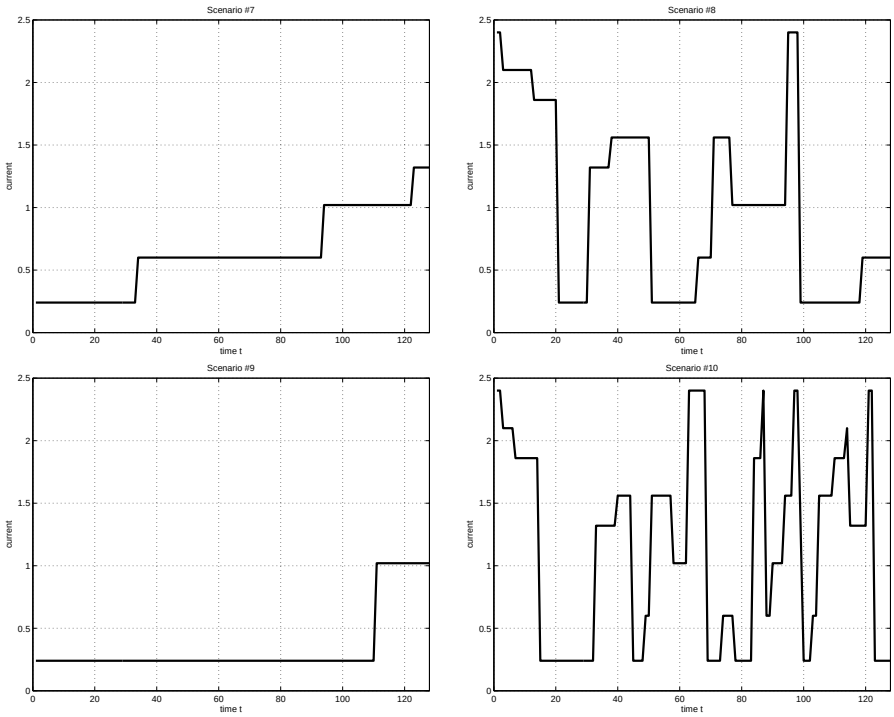

Figure 2. Considered workload scenarios, which are essentially based on the iPhone battery current measurements of [18].

4.2. Discussions and Performance Comparison

Figure 3 is the simulation results of the 10-fold cross-validation, and describes the charge depletion patterns of the battery and the supercapacitors in the HPSS. The exact meaning of the pictures in the figure is as follows: the *i*th picture shows the charge-depleting patterns obtained by the proposed ADP-based DPM procedure for the *i*th numerical experiment, in which the *i*th workload scenario of Figure 2 was used as the test set, and the other nine workload scenarios were used as the training set for estimating the first and second moments of $w(t)$. In the pictures, the solid line represents the charge value of the battery, $r_{bat}(t)$, and the dashed line is for the sum of supercapacitor charges, $r_{cap,1}(t) + r_{cap,2}(t)$. Note that in the sixth picture, the scenario consists of heavy loads (such as playing games, browsing the Internet, and watching YouTube videos), and the charges in the battery and the supercapacitors were completely depleted before the end of the considered time horizon. Figure 3 shows that the proposed ADP-based DPM procedure resulted in reasonable dynamic charge management for the HPSS. From the simulation results, one can see that when the demand charge changed abruptly, then the control actions for the supercapacitors reacted promptly. It is well known that in general, the battery alone cannot handle the situation of rapidly changing loads effectively due to its low power density. By comparison, Figure 3 shows that the HPSS utilizing the ADP-based DPM was capable of handling such situation reasonably well. We believe that this capability is ensured by the stage cost function of Equation (19), which is defined with the purpose of minimizing the battery charge consumption rate and maintaining the charge level of supercapacitors appropriately for speedy reaction to fluctuating loads. Note that our ADP-based DPM procedure can be applied to a significantly larger class of stage cost functions as long as they are quadratic-program (QP) representable [23] (*i.e.*, $l(\cdot,\cdot)$ is convex quadratic plus a convex piecewise linear function, possibly including linear equality and inequality constraints).

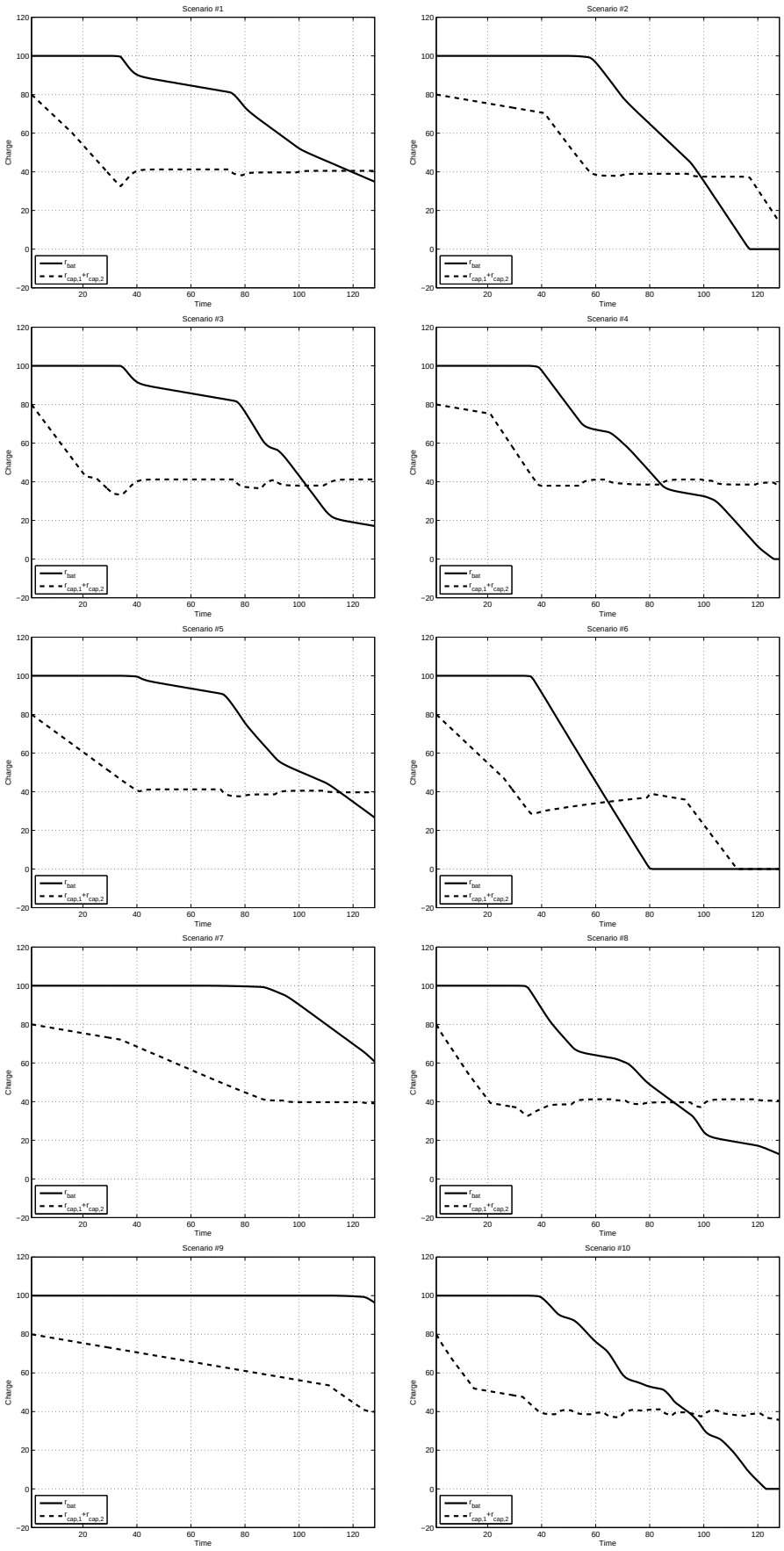

Figure 3. Charge-depleting patterns of the battery and supercapacitors resulting from the ADP-based approach of this paper.

For the sake of performance comparison, we considered the reinforcement learning (RL) based DPM technique of the reference paper [18]. RL is a branch of machine learning which is related to stochastic optimal control. Recently, RL methods [20] have been successfully applied to various decision-making problems, including the field of DPM for HPSSs [18]. One of the main differences of the RL methods for stochastic optimal control is that in their solution procedures, they use the training data on the state transitions together with function approximation instead of exact mathematical analysis. In the reference paper [18], Q-learning [24], which is one of the most widely used value-function-based RL methods, was utilized to derive an RL-based DPM strategy. In Q-learning, given the state transition data ($x(t)$, $u(t)$, $r(t)$, $x(t + 1)$), where $r(t)$ is the reward signal defined as the negative cost (*i.e.*, $r(t) = -l(x(t), u(t))$), the state-action value function Q [20,24] is updated via $Q(x(t), u(t)) \leftarrow Q(x(t), u(t)) + \alpha[r(x(t), u(t)) + \gamma \max_u Q\left(x\left(t+1\right), u\right) - Q(x(t), u(t))]$ In our numerical experiments for performance comparison, the discount rate of Q-learning was set to be the same (*i.e.*, $\gamma = 0.99$). Function approximation for state-action value function Q was performed via the tile coding [20]. Also for the learning rate and the ε value for exploration via the ε-greedy policy [20], we used $\alpha = 0.1$ and $\varepsilon = 0.1$ (For detailed meaning of α and ε in Q-learning, the reader is referred to [20]). Figure 4 shows the learning curve (the solid line) resulting from Q-learning, which describes the average total cost *vs.* the iteration over a set of 10 simulation runs, together with the total cost of the proposed ADP-based DPM approach (the dashed line). As shown in the figure, Q-learning took about 500 iterations of training scenarios to have a convergent behavior. Figure 5 shows the simulation results of the Q-learning-based DPM strategy, and describes the charge depletion patterns of the battery and the supercapacitors in the HPSS. Also, Figure 6 reports the average stage cost sums over a set of ten simulation runs for the two methods. As shown in Figures 5 and 6, the proposed ADP-based DPM strategy yielded better performance than the Q-learning-based DPM [18].

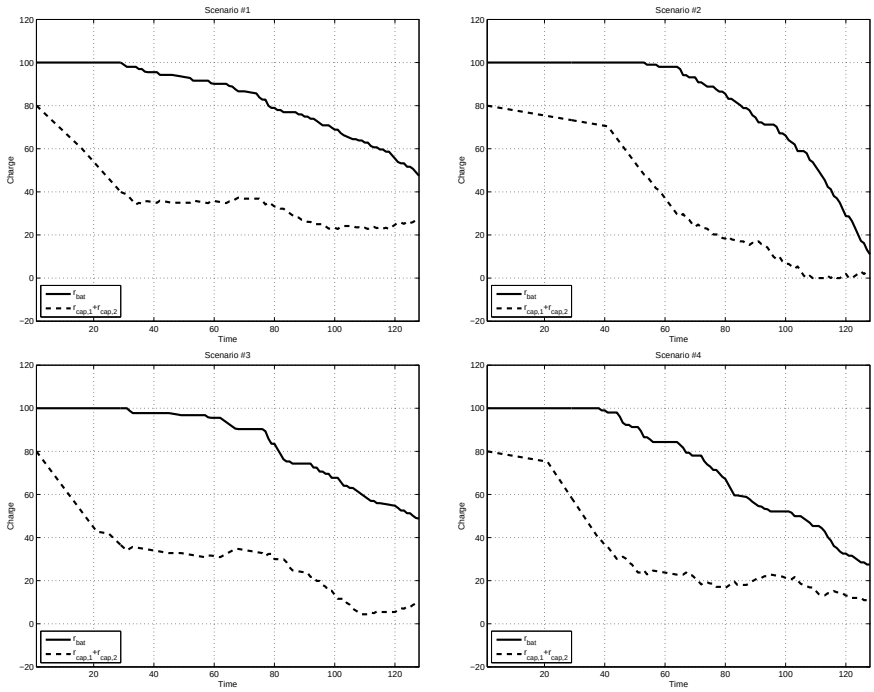

Figure 4. *Cont.*

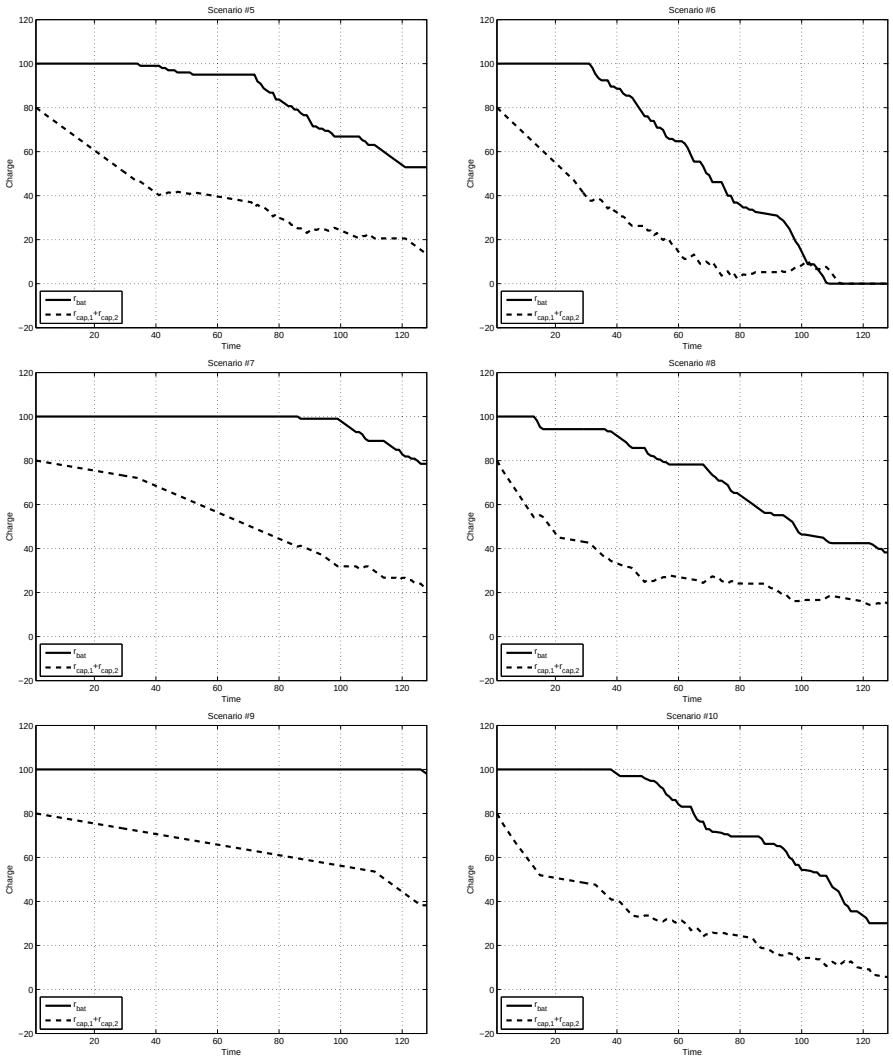

Figure 4. Charge-depleting patterns of the battery and supercapacitors resulting from the Q-learning-based DPM.

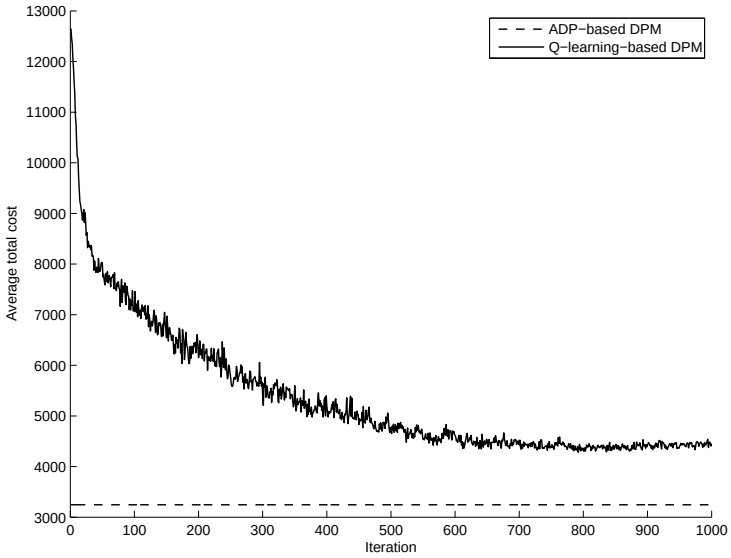

Figure 5. Learning curve (the average total cost *vs.* the iteration over a set of 10 simulation runs): Q-learning (dashed) and ADP (solid).

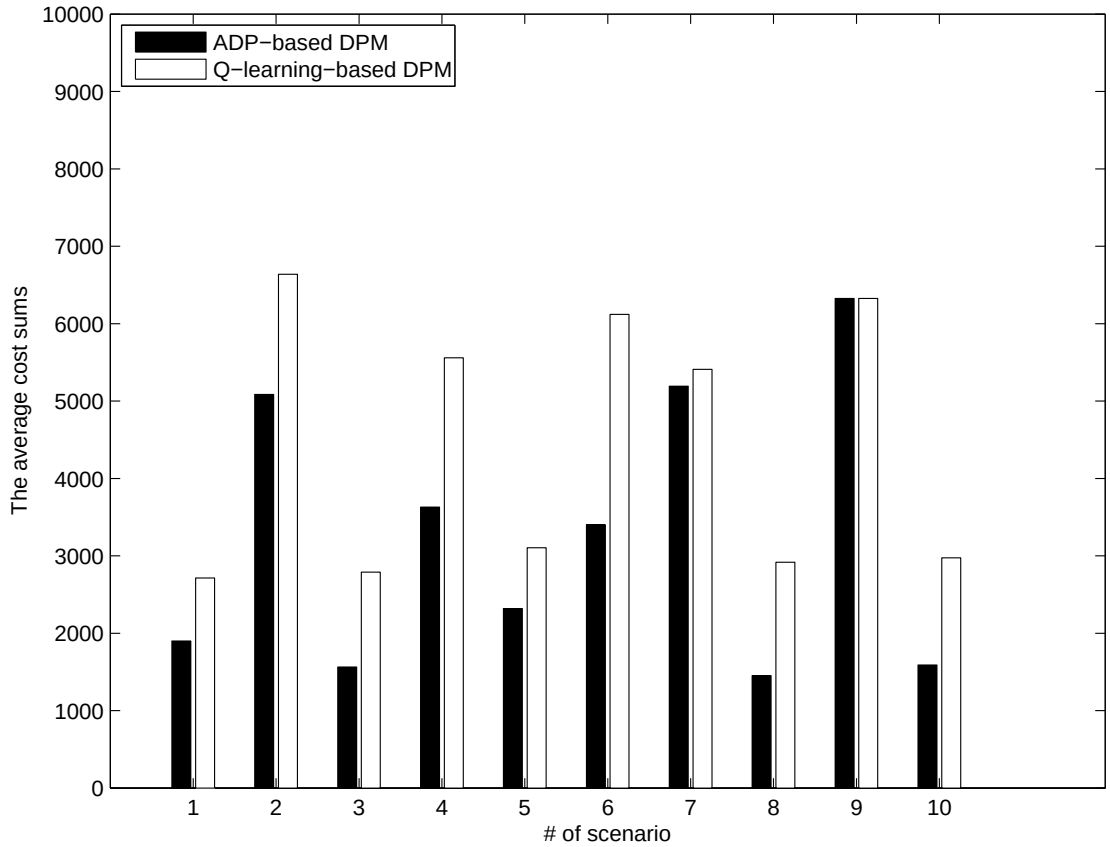

Figure 6. Performance comparison for ADP and Q: Average total cost for each scenario.

Finally, we think that the main significance of the ADP-based DPM approach of this paper and the Q-learning-based approach is in their capacity for customization. More specifically, they can customize the DPM for a (similar) user group, affecting the dynamic power management policy based on the first and second moments of the workloads and/or the past workload scenarios. We expect that this aspect will be covered in future works with sufficiently large training data sets.

5. Conclusions

Managing power flows in portable HPSSs has become an important issue with the advent of a variety of mobile systems and hybrid energy technologies, particularly when supercapacitors are used, which can compensate for the various weaknesses of conventional battery-based power supplies. The problem of optimally managing the power flows for HPSSs thus becomes more complicated. In this paper, the DPM problem for portable HPSSs with combined supercapacitors and battery energy supplies was considered, and an ADP-based DPM strategy was presented. Simulation results show that the proposed ADP-based approach can find a DPM strategy that copes with uncertain workload demands well, and its performance is better compared to the conventional Q-learning-based DPM method. Our plans for future work include more detailed modeling. For example, in the current model of the paper, the losses in the storage systems and in the power converters were neglected, and we are planning to take into account such losses as well with a more detailed model. Also, we will perform more extensive simulation studies, which should reveal the strengths and weaknesses of ADP-based methods, and applications to other types of HPSSs.

Acknowledgments

This work was supported by a Korea University Grant.

Author Contributions

J.P. and G.C. conceived and designed the methodology of the paper; J.P., D.Y. and J.L. wrote the computer program and performed the numerical experiments; J.P. and G.C. wrote the paper.

Conflict of Interest

The authors declare no conflicts of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Burke, A. Ultracapacitors: Why, How, and Where is the Technology. *J. Power Sources* **2000**, *91*, 37–50.
2. Yap, H.T.; Schofield, N.; Bingham, C.M. Hybrid energy/power sources for electric vehicle traction systems. In Proceedings of the IEEE Power Electronics, Machines and Drives (PEMD) Conference, Edinburgh, UK, 31 March–2 April 2004; Volume 1, pp. 61–66.
3. Miller, J.M.; Deshpande, U.; Dougherty, T.J.; Bohn, T. Power electronic enabled active hybrid energy storage system and its economic viability. In Proceedings of the IEEE Applied Power Electronics Conference and Exposition, Washington, DC, USA, 15–19 February 2009; pp. 190–198.
4. Wang, L.; Liu, X.; Li, H.; Im, W.S.; Kim, J.M. Power electronics enabled energy management for energy storage with extended cycle life and improved fuel economy in a PHEV. In Proceedings of the IEEE Energy Conversion Congress and Exposition (ECCE), Atlanta, GA, USA, 12–16 September 2010; pp. 3917–3922.
5. Wang, Y.; Boyd, S. Approximate dynamic programming via iterated bellman inequalities. *Int. J. Robust Nonlinear Control* **2015**, *25*, 1472–1496.
6. Bertsekas, D. *Dynamic Programming and Optimal Control: Volume 1*; Athena Scientific: Belmont, MA, USA, 2005.
7. Bertsekas, D. *Dynamic Programming and Optimal Control: Volume 2*; Athena Scientific: Belmont, MA, USA, 2007.
8. O'Donoghue, D.; Wang, Y.; Boyd, S. Min-Max approximate dynamic programming. In Proceedings of the 2011 IEEE International Symposium on Computer-Aided Control System Design (CACSD), Denver, CO, USA, 20–23 September 2011; pp. 424–431.
9. O'Donoghue, B.; Wang, Y.; Boyd, S. Iterated approximate value functions. In Proceedings of the European Control Conference (ECC), Zurich, Switzerland, 17–19 July 2013; pp. 3882–3888.
10. Romaus, C.; Bocker, J.; Witting, K.; Seifried, A. Optimal energy management for a hybrid energy storage system combining batteries and double layer capacitors. In Proceedings of the Energy Conversion Congress and Exposition (ECCE), San Jose, CA, USA, 20–24 September 2009.

11. Chen, G.; Bao, Z.J.; Yang, Q.; Yan, W.J. Scheduling strategy of hybrid energy storage system for smoothing the output power of wind farm. In Proceedings of the IEEE Control and Automation (ICCA), Hangzhou, China, 12–14 June 2013; pp. 1874–1878.
12. Choi, M.E.; Kim, S.W.; Seo, S.W. Energy Management Optimization in a Battery/Supercapacitor Hybrid Energy Storage System. *IEEE Trans. Smart Grid* **2012**, *3*, 463–472.
13. Gee, A.M.; Robinson, F.V.P.; Dunn, R.W. Analysis of battery lifetime extension in a small-scale wind-energy system using supercapacitors. *IEEE Trans. Energy Convers.* **2013**, *28*, 24–33.
14. Blanes, J.M.; Gutierrez, R.; Garrigos, A.; Lizan, J.L.; Cuadrado, J.M. Electric vehicle battery life extension using ultracapacitors and an FPGA controlled interleaved buck-boost converter. *IEEE Trans. Power Electron.* **2013**, *28*, 5940–5948.
15. Min, S.W.; Kim, S.J. Optimized installation and operations of battery energy storage system and electric double layer capacitor modules for renewable energy based intermittent generation. *J. Electr. Eng. Technol.* **2013**, *8*, 238–243.
16. Gao, L.; Dougal, R.A.; Shengyi, L. Power enhancement of an actively controlled battery-ultracapacitor hybrid. *IEEE Trans. Power Electron.* **2005**, *20*, 236–243.
17. Mirzaei, A.; Farzanehfard, H.; Adib, E.; Jusoh, A.; Salam, Z. A fully soft switched two quadrant bidirectional soft switching converter for ultra capacitor interface circuits. *J. Power Electron.* **2011**, *11*, 1–9.
18. Mirhoseini, A.; Koushanfar, F. Learning to manage combined energy supply systems. In Proceedings of the IEEE/ACM International Symposium on Low-Power Electronics and Design, Fukuoka, Japan, 1–3 August 2011; pp. 229–234.
19. Plett, G.L. Extended Kalman Filtering for Battery Management Systems of LiPB-based HEV Battery pack: Part 1. Background. *J. Power Sources* **2004**, *134*, 277–292.
20. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 1998.
21. Boyd, S.; Ghaoui, L.E.; Feron, E.; Balakrishman, V. *Linear Matrix Inequalities in System and Control Theory*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1994.
22. Mirhoseini, A.; Koushanfar, F. HypoEnergy: Hybrid supercapacitor-battery power-supply optimization for energy efficiency. In Proceedings of the Design, Automation & Test in Europe Conference & Exhibition, Grenoble, France, 14–18 March 2012; pp. 1–4.
23. Boyd, S.; Mueller, M.T.; O'Donoghue, D.; Wang, Y. Performance bounds and suboptimal policies for multi-period investment. *Found. Trends Optim.* **2014**, *1*, 1–69.
24. Watkins, C. Learning from Delayed Rewards. Ph.D. Thesis, Cambridge University, Cambridge, UK, 1989.