Adaptive optimal control for large-scale nonlinear systems

Michailidis, Lakovos; Baldi, Simone; Kosmatopoulos, Elias B.; Ioannou, Petros A.

DOI
10.1109/TAC.2017.2684458

Publication date
2017

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Automatic Control

An Adaptive Learning-based Approach for Nearly-Optimal Dynamic Charging of Electric Vehicle Fleets

Christos Korkas, Simone Baldi, Shuai Yuan and Elias Kosmatopoulos

***Abstract*—Managing grid-connected charging stations for fleets of electric-vehicles leads to an optimal control problem where user preferences must be met with minimum energy costs (e.g. by exploiting lower electricity prices through the day, renewable energy production, stored energy of parked vehicles). Instead of state-of-the-art charging scheduling based on open-loop strategies that explicitly depend on initial operating conditions, this paper proposes an approximate dynamic programming feedback-based optimization method with continuous state space and action space, where the feedback action guarantees uniformity with respect to initial operating conditions, while price variations in the electricity and available solar energy are handled automatically in the optimization. The resulting control action is a multi-modal feedback which is shown to handle a wide range of operating regimes, via a set of controllers whose action that can be activated or deactivated depending on availability of solar energy and pricing model. Extensive simulations via a charging test case demonstrate the effectiveness of the approach.**

***Index Terms*—Electric Vehicles, Charging Optimization, Approximate Dynamic Programming**

I. Introduction

UPCOMING deployment of plug-in hybrid and fully electric vehicles (EVs) requires the integration of a huge amount of electrical storage into the electric utility grid. The introduction of EVs can not only drastically modify the overall load profile [1], but also introduce uncertainty in the grid since, with the Vehicle-to-Grid (V2G) functionality [2], EVs can also provide energy to the power grid by discharging the battery. Several studies in literature have focused on assessment of techno-economic potential of integrated EV-grid systems [3], on ancillary services that EVs can provide to the smart grid [4] and smart mobility [5] scenarios, on renewable energy sources penetration based on EV proliferation [6], and on cybersecurity architectures for smart EV charging [7]. All these studies agree on one crucial point: developing appropriate algorithms to control the charging/discharging process is of fundamental importance for the widespread diffusion of EVs. EV batteries has been modeled individually [4], [8], [9], or as a single aggregate battery with a single state-of-charge [10]–[12]. Regardless of individual or aggregated EV battery models, managing grid-connected charging stations leads to an optimal control problem where not only user preferences must be met (desired battery state of charge) but energy costs should be minimized, possibly taking into account price variations in the electricity price, available solar energy and stochastic vehicle arrival/departure schedule. This optimal control problem is of difficult solution, mainly due to: stochastic charging/discharging dynamics arising from behavior in arrival and departure time [13]; and the need for the charging algorithm to work under different pricing schemes and availability of renewable energy [14]. Several approaches have been proposed in literature for charging optimization of individual EVs or EV fleets: an overview is given below, and some open problems are identified.

C. Korkas and E. Kosmatopoulos are with the Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, 67100 Greece and Informatics & Telematics Institute, Center for Research and Technology Hellas (ITI-CERTH), Thessaloniki 57001, Greece (email: ckorkas@ee.duth.gr and kosmatop@iti.gr).

S. Baldi and S. Yuan are with Delft Center for Systems and Control, Delft University of Technology, Delft 2628CD, The Netherlands (e-mail: s.baldi@tudelft.nl)

A. Related Work

Roughly speaking, we can divide optimization-based EV charging strategies into open-loop strategies (the control is a time-dependent scheduling profile calculated based on predictable operation of the system) and closed-loop strategies (based on feedback measurements): in the first family, [15] proposes a model predictive control approach with statistical EV arrivals and reduced computational complexity, while [16] proposes an EV classification scheme based on mixed-integer programming for a photovoltaic-powered charging station to reduce the cost of energy trading. The authors in [17] formulate the charging problem as an open-loop cost minimizing problem and an open-loop profit maximizing one. In [18] an event-triggered scheduling scheme for V2G operation is proposed that runs every time an EV connects or disconnects and in [19] an improved version of particle swarm optimization is used for optimal charging. Most of these approaches use a finite-horizon approach: this means that computational complexity is a crucial aspect, since the open-loop optimization routine has to run continuously, usually in a receding-horizon fashion. Thus, one has to look for low complexity solutions either in terms of models (e.g. aggregate battery models) or in terms of decision variables (e.g. short planning horizons): as a consequence approximations or decompositions must be carried out so as to make the problem tractable. For example: in [20] linearization techniques are applied to reduce a mixed-integer nonlinear programming model to a mixed-integer linear programming: in [21] the charging problem is split into hierarchical subproblems to better handle complexity; in [22] a two-stage energy exchange scheduling strategy is

presented where at the first stage the electricity cost of a microgrid is minimized and at the second stage the aggregate charging/discharging power is allocated to each EV. The open-loop nature of these strategies arises from the formulation of the charging problem as an optimal control problem 'a la Pontryagin', involving open-loop control candidates. The advantage of these approaches is that they are 'objective-driven', i.e. the optimization takes care of minimizing a certain operation objective (cost): unfortunately, it is well-known that optimal open-loop controls are fragile [23], e.g. they require to recalculate the trajectories for any initial conditions, and thus they are not suited to work robustly for different conditions.

Controls which are uniform with respect to initial conditions take the form of feedback laws, leading to the aforementioned second family of optimization-based charging strategies, which are designed to achieve and maintain the desired operating conditions by comparing them with the actual operating conditions. In the field of smart energy and EV charging, most feedback laws are based on artificial intelligence: the authors in [14] consider a fuzzy logic-based autonomous controller for EV charging, while in [24] an evolutionary learning framework is developed for dynamic energy management of a smart microgrid, and in [25] a neurofuzzy controller is used for frequency regulation in microgrids with fuel cells. Feedback solutions are not free of challenges: the main problem, as compared to open-loop strategies, is that, since the current operating conditions must be compared with the desired operating conditions, one has to define the desired operating conditions. This typically means replacing the 'objective-driven' approach with a 'rule-driven' approach, which requires a good engineering effort in devising appropriate rules that help the system to perform as desired. See for example the hierarchical intelligent system in [26] used to classify abnormal behaviors in the power system possibly caused by abnormal EVs and other loads, or the index-based approach in [27] used to determine the charging priority of EVs based on surplus power. In [28] EVs are divided into responsive to the pricing signals, and unresponsive EVs that define their charging schedule regardless the cost: finally, most valley filling approaches involve a good deal of appropriately designed rules [29].

In common practice rule-driven approaches are often preferred to open-loop ones, mainly because they can operate over different initial conditions. On the other hand, rule-based approaches lead to acceptable performance only if a good set of rules are selected, which requires a lot of hands-on experience, trial-and-errors, and engineering intuition. From this overview the following relevant question arises: is it possible to combine the objective-driven approach of open loop-based solutions with the robustness of feedback rule-driven ones? We argue that the best of these worlds can be achieved if we manage to embed the charging problem into a multi-modal feedback-based optimization problem. This is made possible by considering the charging problem as the optimal solution of a control problem 'a la Bellman', i.e. involving closed-loop control candidates as explained in the following.

B. Main contribution and approach

The aim of this work is to tackle the problem of intelligent charging/discharging of EVs via a nearly-optimal control approach. An Approximate Dynamic Programming (ADP) method is used. The main motivation for using a Dynamic Programming (DP)-based method is that it leads to feedback-based optimal results (see [30], [31] and references therein). However, we distinguish ourselves from current literature in EV charging/discharging in the following threefold sense: first, while a classical DP algorithm aims at solving the optimal control problem exactly and using a look-up table for the value function, the proposed ADP algorithm aims at solving the optimal control problem iteratively by parameterizing the value function; second, differently from classical ADP approaches based on discrete state space and action space [32], we aim at developing a method with continuous state space and action space (which better represent the continuous charge and power variables involved in the charging problem); third, since a single feedback action is not enough due to different pricing schemes and availability of renewable energy, we embed multiple feedback actions in the optimization problem. The proposed approach will therefore be referred to as Multi-Modal ADP (MM-ADP).

A charging station case study shows that the proposed MM-ADP approach results in a set of controllers whose action is activated or deactivated depending on availability of solar energy and pricing model, thus managing a wide range of operating regimes like rule-driven approaches. At the same time, the activation and deactivation of the feedback action comes out of the optimization, thus recovering the 'objective-driven' feature of open-loop strategies. Comparisons with an open-loop strategy are presented.

The paper is organized as follows: Section II describes the EV fleet model, the case study and the different operating conditions. Section III presents the control objectives, as well as some charging approaches used for comparison. Section IV presents the proposed MM-ADP algorithm and its attributes. Validation results and the robustness evaluation of the proposed charging strategy are presented in Section V.

II. Charging Station settings

In this section the basic characteristics of the charging problem are presented. A grid-connected charging station, shown in Fig. 1, is composed of multiple charging spots (in our numerical study we consider 10 charging spots). Every charging spot can be used more than once throughout the day. It is assumed that the grid can satisfy all EVs demands up to their maximum charger output: this assumption allows the maximum possible grid transactions at every time step. The control objective of the charging station is that each battery state of charge (SoC) is at a certain desirable level just before the vehicles departure time. We set a desirable SoC of 100% at departure time. Furthermore, we consider that when beneficial to the charging cost, the charging station can use the energy stored in EVs whose departure time is far in time to charge EVs that are planning to leave shortly in time. In the following we will give the mathematical description of a set of two

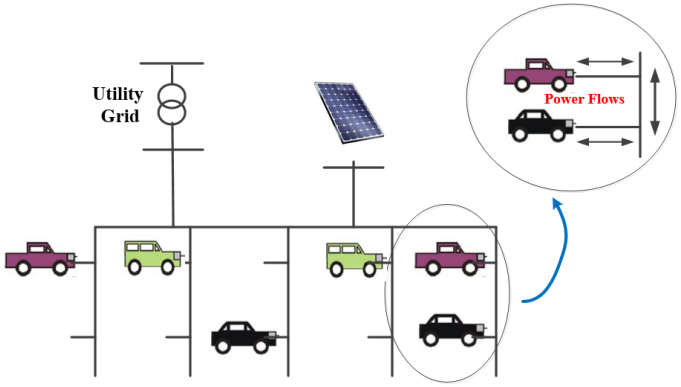


Fig. 1: The charging station and energy transactions between EVs and the grid.

models used to describe the charging problem: the first model is used for simulation and validation and operates at the single batter level), while the second model is used for control design and operates at the aggregate battery level).

TABLE I: Charging station parameters.

Constant parameter	Value
Interval length (dt) (h)	1
Battery Capacity (B_{max}) (kWh)	20
Charging and Discharging Eff. (η_{ch}) (%)	91
Charger Output ($P_{ch,max}$) (kW)	11
Intervals before leaving (k_{leave}) (-)	2

Stochastic parameter	Minimum	Maximum
Arrival State of Charge (%)	10	80
Arrival Time (hour)	0	22
Departure Time (hour)	Arrival+2	Next Day

In Table I, the basic parameters of the EVs and of the charging station are presented. We assume that the fleet of EVs is composed of the same type of vehicles with the same battery capacity and charging/discharging efficiency. The second part of Table I shows that each the arrival SoC, arrival and departure time of each single vehicle are determined by a Gaussian distribution $\mathcal{N}(\mu,\sigma^2)$ with range limited to $\pm 3\sigma$, where μ is the mean, σ^2 is the variance, and the resulting range is indicated in Table I. In the next subsection we describe how to aggregate the single batteries, leading to a stochastic model describing the EV aggregate battery cycle. The advantage of considering Gaussian distributions, e.g. as in [14], [33], [34], is that the Dynamic Programming framework is well defined in this stochastic setting.

A. Fleet Model

The EV battery cycle (charging/discharging) is studied during a period of time evenly divided into intervals of length dt (in this work we consider intervals of 1 hour). We assume that the charging or discharging power within an interval is constant, with new arrivals and departures occurring at the end of the interval, but not inside it. Every interval, N_{plug} are plugged in the station. We consider two separate categories of vehicles:

- N_{leave}, the number of EVs that are scheduled to leave within the next k_{leave} intervals;
- $N_{stay} = N_{plug} - N_{leave}$, the number of EVs that are going to stay during the next k_{leave} intervals.

where k_{leave} is a design parameter of the charging station. At time k, each EV i has a different state of charge $SoC_i(k)$. For every period k, based on N_{leave} and on the charging efficiency η_{ch}, the aggregate power demand and the average power demand in the next k_{leave} intervals can be calculated as:

$$P_{dem,agg}(k) = \frac{\sum_{i=1}^{N_{leave}} ((1-SoC_i(k))B_{max})}{\mathrm{d}t\, k_{leave}\, \eta_{ch}} \quad [kW] \tag{1}$$

which means that the N_{leave} vehicles are charged till they reach full charge in k_{leave} steps. Furthermore, based on N_{stay} and on the discharging efficiency (which is taken equal to the changing efficiency) the aggregate power stored in the staying EVs can be calculated as:

$$P_{stored,agg}(k) = \frac{\sum_{i=1}^{N_{stay}} (SoC_i(k)B_{max})}{\mathrm{d}t\, \eta_{ch}} \quad [kW] \tag{2}$$

Thus, for every interval, the charging control algorithm should meet the requirements $P_{dem,agg}$ by utilizing power P_s from the solar panel when available, power $P_{stored,agg}$ from the staying vehicles whenever convenient, or power P_{grid} from the grid otherwise.

Because of the limited charger output $P_{ch,max}$, the following constraints should be met:

$$P_{dem,agg}(k) \leq N_{leave} P_{ch,max}$$
$$P_{grid}(k) \leq N_{stay} P_{ch,max}$$

Also the battery of each EV should satisfy the upper and lower SoC constraints, as given by

$$0 \leq SoC_{min} \leq SoC_i(k) \leq SoC_{max} \leq 1$$

Summarizing, a stochastic difference equation can be defined for the overall fleet as in (3)-(6). In (3)-(6), we have that SoC_{stay} is the aggregate state of charge of staying vehicles, $P_{s,ex}$ is the excess of solar power, P_s is modeled as a stochastic harmonic oscillator with states $\bar{\bar{P}}_s$, $\bar{P}_s$ and frequency $\omega = 24$ hours $= 86400$ seconds. The parameter $P_{s,avg}$ is the mean production of solar power. In addition, the term S_{stay} can be

negative (in which case the energy from the staying EVs is used for the leaving EVs), or positive (standard charging mode where the power is absorbed from the PV panel or from the grid). Furthermore, the stochastic noise ξ in (3) is due to the stochastic arrival time, stochastic state of charge of the arriving vehicles (collected in SoC_{flow}), and to fluctuations in solar power ΔP_s. The power absorbed from the grid in [kW] is P_{grid}, and p is the electricity pricing in [€/kWh]. Finally, the term π in (6) represents the operating costs at time step k, which includes both the electricity costs and a quadratic term in u to regulate the control authority.

TABLE II: Control design stochastic parameters.

Stochastic Parameter	Mean	Variance
SoC_{flow}	0	0.02
ΔP_s	0	0.001

The following comments on model (3)-(6) are worth underlying:

- The model (3)-(6) is used for control design, while the single-battery models whose parameters are in the second part of Table I is used for simulation and validation purposes. In other words, in the single-battery model, instead of simulating the aggregate SoC_{stay} of staying vehicles, we simulate each single battery.
- On a related note, for control design we model solar power as a stochastic oscillator, while the simulations involve the photovoltaic panel model used in [33]. Please note that the solar power $P_{s,ex}$ in (3) is actually normalized to have the dimensions of a state of charge. The weather data (outside temperature and solar radiation) used to calculate solar production have been downloaded from the EnergyPlus website [34] for the city of Athens, year 2011. The stochastic parameters of the model used for control design are shown in Table II.
- Only the charging/discharging process of the staying vehicles can be controlled, while the demand of leaving vehicles represents an uncontrollable demand. The rationale for this setting is motivated by many practical charging algorithms, e.g. [10], [28].
- The input u has the dimension of a power [kW], and it represents the positive/negative power with which we charge/discharge the staying vehicles.

B. Case study and different conditions

Fig. 2 shows the number of the plugged-in EVs and the $P_{dem,agg}$ throughout one week resulting from the selected stochastic parameters.

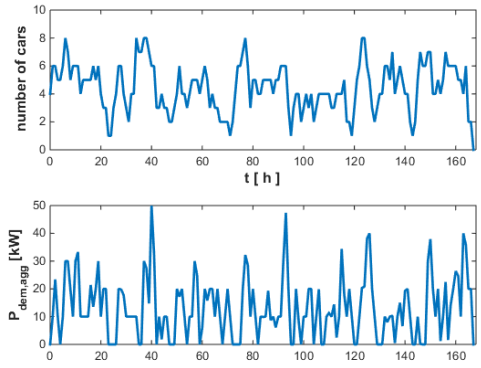


Fig. 2: Evolution of the EV number and $P_{dem,agg}$ for a week.

Since we are interested in how to handle different scenarios we consider both different pricing settings and different average solar energy production settings (represented by different p and $P_{s,avg}$ in (3)-(6)).

For the different pricing scenarios, four different pricing models are selected, shown in Fig. 3. Pricing models 1 and 2 have been obtained from [10], while pricing model 3 is taken from [35], with more complex pricing levels than the previous two models. Finally, the model 4 taken from [36] considers late night hours as high demand hours. Having different pricing models implies having different $p(k)$ in the cost (6).

For the different renewable energy production, three different mean solar energy production are selected: high energy production, medium energy production, and low energy production, which implies having different $P_{s,avg}$ in (4). It is to be expected that different pricing models and different energy production call for different charging/discharging strategies, according to the control objectives formalized in the next section.

III. Control Objectives

One objective of the optimal charging program is to reduce the operating costs (6). Since the pricing models, solar

$$\underbrace{\begin{bmatrix} SoC_{stay}(k+1) \\ S_{stay}(k+1) \\ \bar{\bar{P}}_s(k+1) \\ \bar{P}_s(k+1) \end{bmatrix}}_{x(k+1)} = \underbrace{\begin{bmatrix} SoC_{stay}(k)+S_{stay}(k)+P_{s,ex}(k) \\ S_{stay}(k) \\ \bar{P}_s(k) \\ (2-\omega^2)\bar{P}_s(k)-\bar{\bar{P}}_s(k) \end{bmatrix}}_{f(x(k))} + \underbrace{\begin{bmatrix} 0 \\ \frac{\mathrm{d}t}{B_{max}} \\ 0 \\ 0 \end{bmatrix}}_{B} u(k) + \underbrace{\begin{bmatrix} SoC_{flow}(k) \\ 0 \\ 0 \\ \Delta P_s(k) \end{bmatrix}}_{\xi(k)} \tag{3}$$

$$P_s(k) = \bar{P}_s(k) + P_{s,avg}, \qquad P_{s,ex}(k) = \min[0, P_s(k) - \frac{P_{dem,agg}(k) B_{max}}{\mathrm{d}t k_{leave} \eta_{ch}}] \tag{4}$$

$$P_{grid}(k) = \min[0, P_{dem,agg}(k) + S_{stay}(k) - P_s(k)] \tag{5}$$

$$\pi(x(k), u(k)) = \mathrm{d}t\, p(k) P_{grid}(k) + \rho u^2(k), \quad \rho > 0 \tag{6}$$

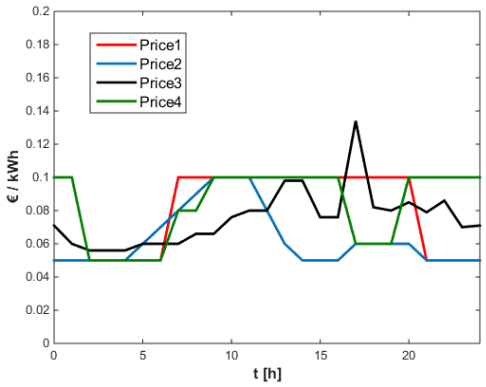


Fig. 3: Evolution of the pricing models used in the case studies.

radiation and EV schedules presented in Section II create a dynamical situation, it is important to consider the effect of the cost over longer horizons,

$$J = \mathscr{E}\left[\pi(k)\right] \approx \frac{1}{N_t}\sum_{k=0}^{N_t-1}\left[\mathrm{d}t\, p(k) P_{grid}(k) + \rho u^2(k)\right] \quad (7)$$

where N_t is the length of the horizon (e.g. N_t = 24 for a day-long experiment or N_t = 168 for week-long experiment), and $\mathscr{E}$ indicates expectation. The average cost (7) will be considered in this work. Minimizing (7) takes into account that if stored and produced energy is not enough, some energy must be absorbed from the main grid at a certain price. Thus, the energy cost can be kept as low as possible by charging the EVs when the electricity price is low and by using stored energy of the aggregate battery and solar panels when electricity price is high. Summarizing, (3)-(6) we obtain the following dynamic optimization problem

$$\begin{aligned} \min\ J &= \mathscr{E}\left[\mathrm{d}t\, p(k) P_{grid}(k) + u^2(k)\right] \quad (8)\\ &\quad s.t. \\ x(k+1) &= f(x(k)) + Bu(k) + \xi(k) \quad (9) \end{aligned}$$

where all the variables have been defined in (3)-(6).

In the following we present three approaches to tackle (8)-(9): the first one is a simple Rule-Based (RB) controller used as the "baseline" charging strategy for comparison purposes. The second one is an Open-Loop Optimization (OLO) which gives a time-dependent scheduling profile for the entire horizon. Finally, in Section IV, we present our proposed Multi-Modal Approximate Dynamic Programming (MM-ADP) algorithm, which implements a truly feedback (closed-loop) action.

A. Rule-Based Approach

The RB strategy follows a similar rationale as [10], [35]: the charging is activated only during time steps when $P_{dem,agg}(k) > 0$. The main goal of RB is to fully charge the EVs which are leaving during the next k_{leave} hours and fully exploit the available solar energy. If the stored energy of the aggregate battery $P_{stored,agg}(k)$ and solar energy $P_s(k)$ are not enough, then RB commands the charging station to absorb electricity from the main grid. Moreover, when $P_s(k)$ is available, the station charges EVs independently whether they are leaving or staying. The RB controller is presented below:

```
if P_dem,agg(k) = 0
   if P_s(k) = 0
      do nothing
   else
      P_dem,agg(k) = P_s(k)
   end
elseif P_dem,agg(k) > 0
      P_grid(k) = P_dem,agg(k) − P_stored,agg(k) − P_s(k)
end
```

B. Open-loop optimization strategy

The RB controller is a simple strategy that does not apply any "intelligent" control action. In order to compare MM-ADP with more intelligent optimization-based strategies, an Open-Loop Optimization (OLO) strategy is implemented. The function *fmincon* (from the Matlab Optimization Toolbox) was used to solve the optimization problem (8)-(9) in an open-loop fashion[1]. More specifically, the *fmincon*-based method optimizes a vector of actions which contains the decisions for a specific experiment-horizon (24 decisions for a day, 168 decisions for a week). These open-loop results, which are explicitly dependent on the EV initial conditions, can be compared with the closed-loop solution as delivered by the proposed approximate dynamic programming algorithm.

IV. Approximate Dynamic Programming Approach

Here we present our proposed Multi-Modal Approximate Dynamic Programming (MM-ADP) optimization method for the EV fleet charging. Using dynamic programming arguments, the optimal solution to (8)-(9) solves the following Hamilton-Jacobi-Bellman equation [37]:

$$\begin{aligned} u^o(x) &= \arg\min_u\ \mathscr{E}\left\{\left(\frac{\partial V^o(x)}{\partial x}\right)'(f(x)+Bu) + \right.\\ &\left. Q(k) + u'(k)u(k) + \mathrm{tr}\left(\Sigma\frac{\partial^2 V^o(x)}{\partial x^2}\right)\right\}, \quad (10) \end{aligned}$$

where V^o is the optimal value function, u^o is the optimal control, and Σ is the variance of ξ. The MM-ADP algorithm parametrizes the optimal solution via

$$\begin{aligned} V^o &= z(x)' P^o z(x) + O(1/L) \quad (11)\\ u^o &= -\rho^{-1} B' M_z(x) P^o z(x) + O(1/L) \quad (12) \end{aligned}$$

where P^o is the parameterization of the optimal value function, with $\varepsilon_1 I \le P^o \le \varepsilon_2 I$, $M_z(x)$ is the Jacobian matrix of $z(x)$ with respect to x, $z(x)$ is the feedback vector of the algorithm, and $O(1/L)$ is the approximation term which becomes smaller by increasing L. The exact form of L and $z(x)$ will be discussed in Section IV.A. It is important to note that the control formulation 'a la Bellman' involves closed-loop control candidates $u^o(x)$ in place of open-loop solutions: thus this

[1]The function *fmincon* was selected in view of the nonlinear min functions contained in (3)-(6).

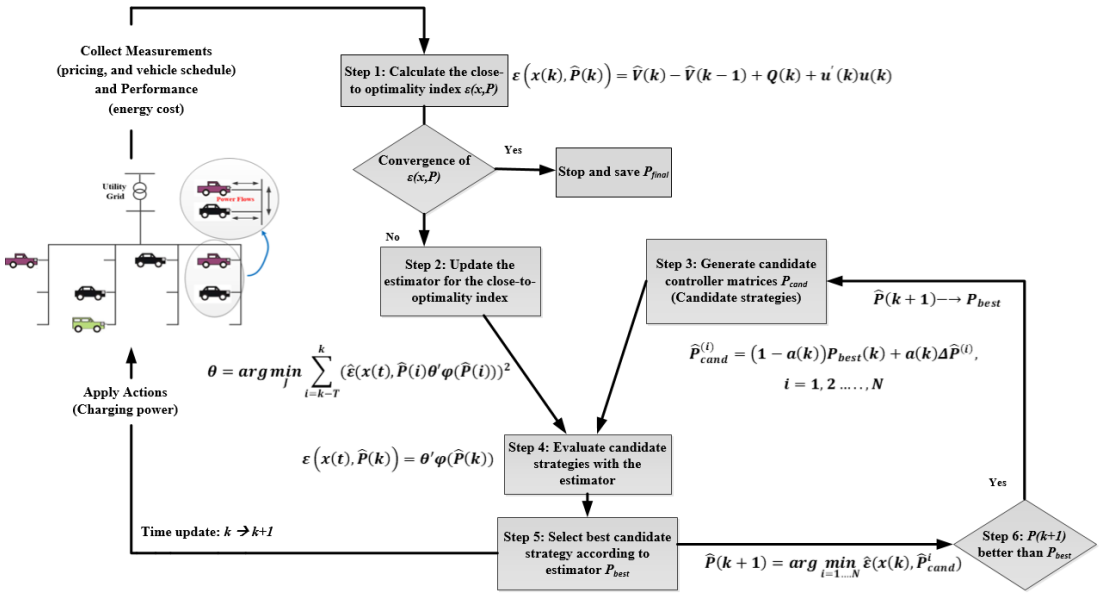


Fig. 4: Flow diagram of the proposed Multi-Modal Approximate Dynamic Programming algorithm.

formulation is expected to give uniform performance with respect to initial conditions.

Since the optimal parametrization P^o is unknown, the idea is to iteratively find the nearly-optimal solution by updating the parametrization at every time step, i.e.

$$\hat{V} = z(x)'\hat{P}(k)z(x) \quad (13)$$
$$\hat{u} = -\rho^{-1}B'M_z(x)\hat{P}(k)z(x) \quad (14)$$

where $\hat{P}$ is an estimate of the P^o, with $\varepsilon_1 I \leq \hat{P}(k) \leq \varepsilon_2 I$. The MM-ADP algorithm modifies/updates the parameters of the controller in order to approach the optimal solution by evaluating them through the use of the charging station model. The MM-ADP algorithm is schematically described in Fig. 4.

- STEP 1 (Calculation of close-to-optimality index): in the first step, the close-to-optimality index

$$\varepsilon(x(k),\hat{P}(k)) = \hat{V}(k)-\hat{V}(k-1)+Q(k)+u'(k)u(k) \quad (15)$$

is used to provide a "measure" of how far the estimate $\hat{P}$ is form its optimal value P^o. The objective is to develop a gradient-like descent method for updating $\hat{P}$,

$$\hat{P}(k+1) = \hat{P}(k) - \gamma\nabla_{\hat{P}}E(\hat{P}(k)), \quad \gamma > 0 \quad (16)$$

in an attempt to minimize the error term $E(\hat{P}) = \varepsilon^2$ and to make $\hat{P}$ converge as close as possible to the nearly-optimal P^o. However, (16) cannot be directly used because of the following three important problems:

1) An exact expression of the gradient $\nabla_{\hat{P}}E$ in (16) is not available, since the gradient depends by the dynamics in (9) which are affected by stochastic noise.
2) Moreover, in (16), the $\hat{P}$ matrix must remain positive definite. Thus, a constrained gradient descent must be implemented, increasing the complexity (due to the use of penalty function, generalized Lagrangian multiplier, etc).
3) The stochastic noise and the approximation term $O(1/L)$ can "destroy" the convergence properties of a standard gradient descent algorithm [38].

To overcome all the above problems, we proceed with the following steps.

- STEP 2 (Update linear-in-the-parameters estimator): The following linear-in-the-parameters estimator,

$$\hat{\varepsilon}(x(k),\hat{P}(k)) = \theta'\phi(\hat{P}(k)) \quad (17)$$
$$\theta = \arg\min_{\vartheta}\sum_{i=k-T}^{k}\left(\hat{\varepsilon}(x(k),\hat{P}(i)) - \vartheta'\phi(\hat{P}(i))\right)^2 \quad (18)$$

is adopted to approximate the gradient of the objective function with respect to $\hat{P}$. As shown in [39], this adaptive and stochastic approximation technique can approximate the performance of gradient descent methods.
- STEP 3 (Generate candidate strategies): only positive definite matrices are considered. This is achieved by generating the following candidate perturbations

$$\hat{P}^{(i)}_{cand} = (1-a(k))\hat{P}_{best}(k) + a(k)\Delta\hat{P}^{(i)}, i = 1,2,\ldots,N \quad (19)$$

where $\Delta\hat{P}^{(i)}$ is a random positive definite matrix, which guarantees $\varepsilon_1 I \leq \hat{P}(k) \leq \varepsilon_2 I$ at every time k.
- STEP 4 (Evaluate strategies) and 5 (Select best strategy): Not all the candidate perturbations are evaluated via a simulation model of the charging station. The candidate perturbations are evaluated by the estimator, and only the best one (according to the estimator) will be evaluated

$$\hat{P}(k+1) = \arg\min_{i=1,\ldots,N}\hat{\varepsilon}(x(k),\hat{P}^{(i)}_{cand}) \quad (20)$$

The use of estimator (17), results in only one evaluation per time step of the objective function via the charging station model.
- STEP 6 (Simulation-based reset): In order to reduce the effect of the term $O(1/L)$, a reset based on the the

simulation-based evaluation is used. In fact, the performance of $\hat{P}(k+1)$ is compared with $\hat{P}_{best}(k)$, which is the value function that gave the best performance till that time.

The convergence properties of the proposed MM-ADP algorithm are summarized by the following theorem.

Theorem 1: The MM-ADP algorithm depicted in Fig. 4, guarantees that $\hat{P}_k$ converges with probability 1 to the set

$$\mathscr{C} = \left\{\hat{P} : \hat{P} \text{ is positive definite and } \nabla_{\hat{P}} \varepsilon^2 \left(x(k), \hat{P}\right) = 0\right\}$$

Proof. The proof is not shown for lack of space, but the interested reader can derive it based on stochastic approximation tools and on the stochastic version of the main result in [40]. Note that, as discussed in [40], convergence to $\mathscr{C}$ might lead to local nearly-optimality, and not necessarily global nearly-optimality.

A. Multi-modal control action

The MM-ADP strategy exploits different information to operate an optimal management: such information is collected in the feedback vector x. The feedback vector is then used to approximate the value function and the control law: in most cases a quadratic approximation of the value function $x'\hat{P}x$ and a linear approximation of the control law $\hat{u} = -\rho^{-1}B'\hat{P}x$ can provide acceptable performance. However, since (3)-(6) that we must handle different pricing models $p(k)$ and different mean solar production $P_{s,avg}$, this might lead to too different charging dynamics to be handled by a single controller. For example, days with high solar radiation call for a completely different charging strategy than days with low solar radiation, while different pricing models calls for completely different charging strategies. Pretending that a single linear control strategy can handle every situations in a nearly-optimal way is not realistic. Therefore, the quadratic/linear approximations must be overcome, and it may be better to utilize different controllers for different operating regimes. In that case, $\hat{P}$ and $z(x)$ can be formulated as follows:

$$\hat{P}(k) = \begin{bmatrix} P_1 & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_L \end{bmatrix}, \quad z(k) = \begin{bmatrix} \sqrt{\beta_1(p, P_{s,avg})}x(k) \\ \sqrt{\beta_2(p, P_{s,avg})}x(k) \\ \cdots \\ \sqrt{\beta_L(p, P_{s,avg})}x(k) \end{bmatrix}$$

where L stands for the number of modes we want to consider. The above approximations express the ability of using L different controllers, one for each appropriate case. The activation of each controller determined by β_i, $i = 1...L$, where $\beta_i = 1$ or $\beta_i = 0$ depending on the active p and $P_{s,avg}$. Rather than 'rule-driven' as in standard artificial intelligence (e.g. fuzzy) [41], [42] the resulting multi-modal control action is objective-driven, according to the minimization of (8)-(9).

V. Results

This section is organized as follows: for $L = 1$ (one mode) MM-ADP is compared with RB and OLO without solar energy and later with solar energy; then for $L > 1$ the benefits of multi-modal control is tested in different conditions. In our charging test case we consider two different choices for L, i.e. $L := L_p$ modes corresponding to 4 pricing models and $L := L_s$ modes corresponding to 3 mean solar production. The irradiance data used to calculate solar production are taken from the EnergyPlus database for the city of Athens during the winter (low solar energy production) and summer (high solar energy production) of 2011 [34]. All algorithms are run on the following platform: a PC using 16GB of RAM and an Intel 4770k CPU. Furthermore, It is important to underline that both OLO and MM-ADP are run for a maximum of 1 hour in order to reflect the real-time constraint of having a solution within one time step.

A. Validation results without solar energy

This first subsection is focused on simulations without the presence of solar power with a unique mode for all pricing models. This means that MM-ADP will activate a unique feedback controller. Our aim is not only to showcase the ability of the proposed algorithm to reduce the cost even in cases of zero solar production. Table III presents the charging cost in € for the four pricing models presented in Section II: note that, the costs are averaged over seven 1-day simulations and seven 1-week simulations, in order to take into account the stochastic characteristics of the charging problem.

TABLE III: Average cost in € for 1-day and 1-week simulations (the average is done over seven realizations).

Day	Price 1	Price 2	Price 3	Price 4
RB	23.89	17.07	20.08	21.65
OLO	19.99	14.42	17.29	16.95
MM-ADP	18.71	14.95	17.35	16.90
Week	Price 1	Price 2	Price 3	Price 4
RB	151.46	118.78	141.60	155.25
OLO	140	105.32	130.56	132.44
MM-ADP	128.09	100.45	124.45	121.05

For 1-day experiments, MM-ADP attains approximately 5€ savings over RB, whereas for 1-week experiments, it attains 25-30€ savings. The improvement is in the range $12-22\%$ depending on the pricing model. Also, the MM-ADP algorithm provides solutions that are better, in most cases, than the OLO approach (especially for 1-week experiments): the reason lies in the open-loop nature of the *fmincon*-based algorithm. Since the same 24 (for one day) or 168 (for 1 week) control actions will be repeated over all 7 realization, this set of actions cannot be optimal for all initial conditions for EVs schedules: on the other hand, the feedback nature of the MM-ADP control strategy is independent from these initial conditions. These experiments validate the fact that MM-ADP performance is uniform over a wide range of initial conditions.

B. Validation Results with Solar Energy

This subsection is focused in simulations in the presence of solar power, and also involves a unique mode for MM-ADP. Our aim is to showcase the effect of available solar production in reduction of charging cost.

Table IV presents, for high solar energy production (summer data) the charging cost of each algorithm, for the four pricing

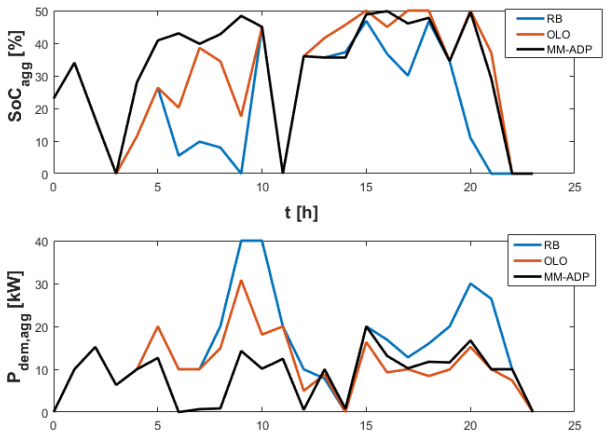


Fig. 5: Evolution of SoC_{agg} (for staying EVs, upper plot) and $P_{dem,agg}$ (for leaving EVs, lower plot).

models. Similarly to the results without solar energy, the costs in Table III are averaged over seven for 1-day simulations and seven 1-week simulations.

TABLE IV: Average cost in € for 1 day and 1 week with *high* solar production (the average is done over seven realizations).

Day	Price 1	Price 2	Price 3	Price 4
RB	12.31	10.44	12.28	14.28
OLO	10.48	8.86	11.23	12.28
MM-ADP	9.96	8.75	11.3	11.46

Week	Price 1	Price 2	Price 3	Price 4
RB	84.02	73.96	92.54	93.43
OLO	78.79	69.56	88.23	80.91
MM-ADP	77.03	69.05	88.45	80.34

MM-ADP has again lower charging costs, as compared to the RB controller. For 1-day experiments, it attains approximately 2-3 € savings, whereas for 1-week experiments, it attains 10 €. The improvement is in the range $8-19\%$ depending on the pricing model. The main reason for smaller improvement than Table III, is to be expected. In fact, Fig. 5 shows the evolution of the aggregate *SoC* of the staying EVs and the demand of the leaving EVs through one day (one particular realization), for the RB, OLO and MM-ADP strategies: when solar energy is available, it can satisfy part of the demand (and sometimes solar energy can be in excess so that it can even charge the staying vehicles). Thus, it is intuitive to say that the more the solar energy, the less the room for improvement for any charging algorithm.

In order to confirm this, Table V presents the charging cost of each algorithm with low solar energy production (winter data). For 1-day experiments, MM-ADP attains approximately 3-4 € savings over RB, whereas for 1-week experiments, it attains more than 10-15 €. The improvement is in the range $10-21\%$ depending on the pricing model. As expected, cost savings lie in between the improvements achieved in Table III and Table IV.

TABLE V: Average cost in € for 1 day and 1 week with *low* solar production (the average is done over seven realizations).

Day	Price 1	Price 2	Price 3	Price 4
RB	18.21	14.11	17.2	18.57
OLO	16.5	12.28	15.42	15.46
ADP	15.11	12.42	15.52	14.58

Week	Price 1	Price 2	Price 3	Price 4
RB	120	96.92	114.79	127.55
OLO	112.64	90.25	112.08	115.11
ADP	98.89	87.16	108.72	100.01

C. Validation Results in Diverse Conditions

The previous simulations have shown that, even with a single mode, MM-ADP shows some robustness. In this subsection we would like to show the capability to handle a wide range of operating conditions by exploiting multiple modes ($L > 1$). To this purpose, two different 15-day simulations were created:

- In one 15-day experiment we alternate among the four pricing models (this is done by changing p in (3)-(6), and using $L := L_p = 4$);
- In another 15-day experiment we alternate among three levels of average solar production (this is done by changing $P_{s,avg}$ in (3)-(6) and using $L := L_s = 3$);

In addition, for both experiments we consider not only Gaussian distribution, but also uniform distribution for arrival *SoC*, arrival and departure times. The range for the uniform distribution are the same as in Table I. Similarly to previous results, every simulation is averaged over seven realizations with different EV arrival schedules and *SoC*. Fig. 6 shows six curves for the evolution of the solar energy during one day for low, medium, and high mean solar production. Note that the order in which the pricing models (or average solar energy) are alternated during the 15 days is not important: any order would do. The true challenge arises from improving the cost by using different control action in place of a single one: in addition, the interest of considering uniform distribution is on verifying robustness to unmodelled dynamics. In these simulations OLO is tested in the most extreme case, because it must operate not only under different initial conditions, but also different pricing model and average solar energy.

TABLE VI: Average cost in €, 15-day simulation with different pricing models, for Gaussian and uniform distributions (the average is done over seven realizations).

	RB	OLO	MM-ADP $L_p = 1$	MM-ADP $L_p = 4$
Gaussian	310.92	291.28	269.34	260.76
Uniform	315.54	294.49	273.07	263.96

Table VI, which presents the charging costs for the charging strategies, shows that a further 3% improvement (from 13% to 16%) is achieved for MM-ADP with $L_p = 4$ as compared with MM-ADP with $L_p = 1$, which implies that the MM-ADP optimization recognizes that a single controller cannot handle such diverse pricing models in an optimal way. Similar improvements apply for the uniform distribution case.

Table VII, which presents the charging costs for the charging strategies, shows that a further 4% improvement (from 15% to

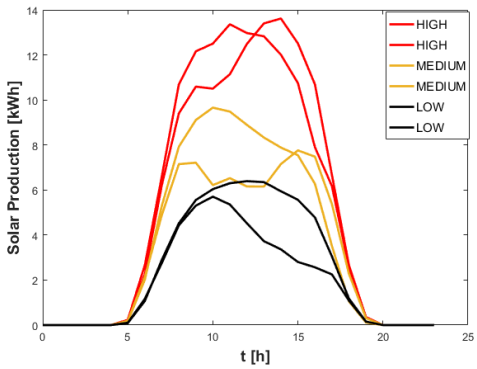


Fig. 6: Daily evolution of the solar energy for low, medium, and high mean solar energy production.

TABLE VII: Average cost in €, 15-day simulation with different solar production, for Gaussian and uniform distributions (the average is done over seven realizations).

	RB	OLO	MM-ADP $L_s = 1$	MM-ADP $L_s = 3$
Gaussian	290.28	282.15	245.94	235.28
Uniform	293.12	286.21	248.34	238.73

19%) is achieved for MM-ADP with $L_s = 3$ as compared with MM-ADP with $L_s = 1$, indicating that the multi-modal optimization enables MM-ADP algorithm to take more focused decisions based on the availability of solar energy of each day.

It would not be too difficult at this point to combine the 4 pricing models and 3 mean solar production to end up with 12 modes corresponding all combinations of operating conditions: simulations are not shown because they would be redundant. Overall, we have shown that MM-ADP is able to manage in a robust and nearly-optimal way a wide range of operating regimes.

VI. Conclusion

This work proposed an intelligent optimization approach based on Multi-Modal Approximate Dynamic Programming (MM-ADP) for the optimal charging/discharging vehicle schedule of a grid-connected charging station. The contribution of this work was: solving the optimal control problem iteratively by parameterizing the value function; considering continuous state space and action space which better represent the continuous charge and power variables involved in the charging problem; embedding multiple feedback actions in the optimization problem. Extensive simulations demonstrated that the proposed strategy exhibits a robust behavior in the presence of stochastic arrival and departure times as well as different pricing models and solar energy production.

Future work will include considering a network of charging stations: this would imply making the proposed algorithm distributed and take into account the power grid constraints in terms of voltage and frequency.

Acknowledgment

The research leading to these results has been partially funded by the European Commission FP7-ICT-2013.3.4, Advanced computing, embedded and control systems, under contract #611538 (LOCAL4GLOBAL).

References

[1] Z. Ma, D. S. Callaway, and I. A. Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *Control Systems Technology, IEEE Transactions on*, 21(1):67–78, 2013.

[2] C. Guille and G. Gross. Design of a conceptual framework for the v2g implementation. In *Energy 2030 Conference, 2008. ENERGY 2008. IEEE*, pages 1–3. IEEE, 2008.

[3] P. Finn, C. Fitzpatrick, and D. Connolly. Demand side management of electric car charging: Benefits for consumer and grid. *Energy*, 42(1):358–363, 2012.

[4] N. Rotering and M. Ilic. Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets. *Power Systems, IEEE Transactions on*, 26(3):1021–1029, 2011.

[5] S. Huang, L. He, Y. Gu, K. Wood, and S. Benjaafar. Design of a mobile charging service for electric vehicles in an urban environment. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):787–798, 2015.

[6] B. Zhu, H. Tazvinga, and X. Xia. Switched model predictive control for energy dispatching of a photovoltaic-diesel-battery hybrid power system. *IEEE Transactions on Control Systems Technology*, 23(3):1229–1236, 2015.

[7] A. C. F. Chan and J. Zhou. A secure, intelligent electric vehicle ecosystem for safe integration with the smart grid. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):3367–3376, 2015.

[8] T. Lan, J. Hu, Q. Kang, C. Si, L. Wang, and Q. Wu. Optimal control of an electric vehicles charging schedule under electricity markets. *Neural Computing and Applications*, 23(7-8):1865–1872, 2013.

[9] S. Han, S. H. Han, and K. Sezaki. Design of an optimal aggregator for vehicle-to-grid regulation service. In *Innovative Smart Grid Technologies (ISGT), 2010*, pages 1–8. IEEE, 2010.

[10] B. Škugor and J. Deur. Dynamic programming-based optimisation of charging an electric vehicle fleet system represented by an aggregate battery model. *Energy*, 92:456–465, 2015.

[11] J. Zheng, X. Wang, K. Men, C. Zhu, and S. Zhu. Aggregation model-based optimization for electric vehicle charging strategy. *Smart Grid, IEEE Transactions on*, 4(2):1058–1066, 2013.

[12] O. Sundström and C. Binding. Flexible charging optimization for electric vehicles considering distribution grid constraints. *Smart Grid, IEEE Transactions on*, 3(1):26–37, 2012.

[13] E. S. Rigas, S. D. Ramchurn, and N. Bassiliades. Managing electric vehicles in the smart grid using artificial intelligence: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):1619–1635, 2015.

[14] S. Faddel, A.A.S. Mohamed, and O. A. Mohammed. Fuzzy logic-based autonomous controller for electric vehicles charging under different conditions in residential distribution systems. *Electric Power Systems Research*, 148:48 – 58, 2017.

[15] W. Tang and Y. J. Zhang. A model predictive control approach for low-complexity electric vehicle charging scheduling: Optimality and scalability. *IEEE Transactions on Power Systems*, 32(2):1050–1063, 2017.

[16] W. Tushar, C. Yuen, S. Huang, D. B. Smith, and H. V. Poor. Cost minimization of charging stations with photovoltaics: An approach with ev classification. *IEEE Transactions on Intelligent Transportation Systems*, 17(1):156–169, 2016.

[17] M. Zhang and J. Chen. The energy management and optimized operation of electric vehicles based on microgrid. *IEEE Transactions on Power Delivery*, 29(3):1427–1435, 2014.

[18] L.i. Jian, Y. Zheng, X. Xiao, and C.C. Chan. Optimal scheduling for vehicle-to-grid operation with stochastic connection of plug-in electric vehicles to smart grid. *Applied Energy*, 146:150 – 161, 2015.

[19] J. Yang, L. He, and S. Fu. An improved pso-based charging strategy of electric vehicles in electrical distribution grid. *Applied Energy*, 128:82–92, 2014.

[20] J. F. Franco, M. J. Rider, and R. Romero. A mixed-integer linear programming model for the electric vehicle charging coordination problem in unbalanced electrical distribution systems. *IEEE Transactions on Smart Grid*, 6(5):2200–2210, 2015.

[21] M. O. Badawy and Y. Sozer. Power flow management of a grid tied pv-battery system for electric vehicles charging. *IEEE Transactions on Industry Applications*, 53(2):1347–1357, 2017.

[22] D. Wang, X. Guan, J. Wu, P. Li, P. Zan, and H. Xu. Integrated energy exchange scheduling for multimicrogrid system with electric vehicles. *IEEE Transactions on Smart Grid*, 7(4):1762–1774, 2016.

[23] M. Bardi and I. Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Modern Birkhauser Classics, 1997.

[24] G. K. Venayagamoorthy, R. K. Sharma, P. K. Gautam, and A. Ahmadi. Dynamic energy management system for a smart microgrid. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1643–1656, 2016.

[25] P. C. Sekhar and S. Mishra. Storage free smart energy management for frequency control in a diesel-pv-fuel cell-based hybrid ac microgrid. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1657–1671, 2016.

[26] Y. Xu, R. Zhang, J. Zhao, Z. Y. Dong, D. Wang, H. Yang, and K. P. Wong. Assessing short-term voltage stability of electric power systems by a hierarchical intelligent system. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1686–1696, 2016.

[27] L. Jian, Y. Zheng, and Z. Shao. High efficient valley-filling strategy for centralized coordinated charging of large-scale electric vehicles. *Applied Energy*, 186, Part 1:46 – 55, 2017.

[28] E. Xydas, C. Marmaras, and L. M. Cipcigan. A multi-agent based scheduling algorithm for adaptive electric vehicles charging. *Applied Energy*, 177:354 – 365, 2016.

[29] G. Benetti, M. Delfanti, T. Facchinetti, D. Falabretti, and M. Merlo. Real-time modeling and control of electric vehicles charging processes. *IEEE Transactions on Smart Grid*, 6(3):1375–1385, 2015.

[30] M. Cipek, B. Škugor, M. Čorić, J. Kasać, and J. Deur. Control variable optimisation for an extended range electric vehicle. *International Journal of Powertrains*, 5(1):30–54, 2016.

[31] R. E. Bellman and S. E. Dreyfus. *Applied dynamic programming*. Princeton university press, 2015.

[32] S. Xie, W. Zhong, K. Xie, R. Yu, and Y. Zhang. Fair energy scheduling for vehicle-to-grid networks using adaptive dynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1697–1707, 2016.

[33] K. Tanaka, A. Yoza, K. Ogimi, A. Yona, T. Senjyu, T. Funabashi, and C.-H. Kim. Optimal operation of dc smart house system by controllable loads based on smart grid topology. *Renewable Energy*, 39(1):132–139, 2012.

[34] U.S. Department of Energy. Energy-plus weather data. https://energyplus.net/weather, 2017.

[35] H. Lund and W. Kempton. Integration of renewable energy into the transport and electricity sectors through v2g. *Energy policy*, 36(9):3578–3587, 2008.

[36] Nord Pool. Nord pool. http://www.nordpoolspot.com/Market-data1/, 2016.

[37] T. Parisini and R. Zoppoli. Neural approximations for infinite horizon optimal control of nonlinear stochastic systems. *IEEE Transactions on Neural Networks*, 9:1388–1408, 1998.

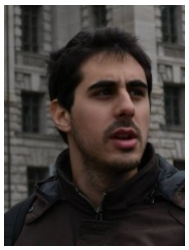
[38] James C. Spall. *Introduction to stochastic search and optimization: estimation, simulation, and control*, volume 65. John Wiley & Sons, 2005.

[39] E. B. Kosmatopoulos and A. Kouvelas. Large scale nonlinear control system fine-tuning through learning. *Neural Networks, IEEE Transactions on*, 20(6):1009–1023, 2009.

[40] S. Baldi, I. Michailidis, E. B. Kosmatopoulos, and P. A. Ioannou. A plug and play computationally efficient approach for control design of large-scale nonlinear systems using cosimulation. *IEEE Control Systems Magazine*, 34:56–71, 2014.

[41] C.-L. Chen and W.-C. Chen. Fuzzy controller design by using neural network techniques. *IEEE Transactions on Fuzzy Systems*, 2(3):235–244, 1994.

[42] J. Y. Choi and J. A. Farrell. Nonlinear adaptive control using networks of piecewise linear approximators. *IEEE Transactions on Neural Networks*, 11(2):390–401, 2000.



Christos D. Korkas graduated from the Department of Electrical and Computer Engineering of Polytechnic School, Democritus University of Thrace in 2013 and currently is a PhD candidate at the same department. His research interests are in the area of adaptive and learning systems, covering applications to smart energy systems. During his research activity he participated in the following three EU FP7 projects: AGILE, LOCAL4GLOBAL, NOPTILUS, the first two involving the development of adaptive energy management solutions for large-scale office buildings.

Simone Baldi received the B.Sc., M.Sc. and Ph.D. degree from University of Florence, Italy, in 2005, 2007, 2011 respectively. He is currently Assistant Professor at the Delft Center for Systems and Control, Delft University of Technology. He held postdoc researcher positions at the University of Cyprus, and at the Information Technologies Institute (I.T.I.-CE.R.T.H). His research interests include adaptive control, switching supervisory control and approximately optimal control, with applications in energy efficient buildings and intelligent transportation systems.

Shuai Yuan received the B.Sc. and M.Sc. degree in Mechanical Science and engineering from Harbin Institute of Technology, Huazhong University of Science and Technology, China, in 2011 and 2014 respectively. He is currently a PhD candidate at the Delft Center for Systems and Control, Delft University of Technology, Delft, the Netherlands. His research interests include adaptive systems and switched systems.

Elias B. Kosmatopoulos received the Diploma, M.Sc. and Ph.D. degrees from the Technical University of Crete, Greece, in 1990, 1992, and 1995, respectively. He is currently an Professor with the Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece. Previously, he was a faculty member of the Department of Production Engineering and Management, Technical University of Crete (TUC), Greece, a Research Assistant Professor with the Department of Electrical Engineering- Systems, University of Southern California (USC) and a Postdoctoral Fellow with the Department of Electrical & Computer Engineering, University of Victoria, B.C., Canada. Dr. Kosmatopoulos research interests are in the areas of neural networks, adaptive optimization and control,energy efficient buildings and smart grids and intelligent transportation systems. He is the author of over 40 journal papers. He has been currently leading many research projects funded by the European Union with a total budget of about 10 Million Euros.