Stochastic DP Based on Trained Database for Sub-optimal Energy Management of Hybrid Electric Vehicles

Rustem Abdrakhmanov(✉) and Lounis Adouane

Institut Pascal/IMobS3, UCA/SIGMA UMR CNRS 6602,
Clermont-Ferrand, France
{rustem.abdrakhmanov,lounis.adouane}@uca.fr

Abstract. This paper presents a sub-optimal energy management strategy, based on Stochastic Dynamic Programming (SDP), for efficient powersplit of a Hybrid Electric Vehicle (HEV). An optimal energy management strategy is proposed, permitting to have simultaneous speed profile and powersplit optimization of the HEV. Formulated as a multi-objective optimization problem, an ϵ-constraint method has been used to find the Pareto front of the energy optimization task. Traffic conditions and driver behavior could be assimilated to a stochastic nature, thus, it is proposed in this paper to address the vehicle power as Markov Decision Process. A Stochastic Database is used to store Transition Probability and Reward Matrices, corresponding to suitable vehicle actions w.r.t. specific states. They are used afterwards to calculate sub-optimal powersplit policy for the vehicle via an infinite-horizon SDP approach. Simulation results demonstrate the effectiveness of the proposed approach compared to a deterministic strategy given in [1]. The present work is conducted on a dedicated high-fidelity model of the HEV that was developed on MATLAB/TruckMaker software.

Keywords: Hybrid electric vehicle · Energy management · Stochastic dynamic programming · Markov decision process · Multi-objective optimization · ϵ-constraint

1 Introduction

In recent years, the problem of reducing the level of pollution by vehicle's exhaust gazes (mainly in urban zone) has become one of the main research topics. This issue has been considered and recognized at the state governments level. Multiple laws and decrees have been passed to smoothly switch over to a zero-emission technologies (e.g., pure electric vehicles (EV)). Hybrid electric vehicle (HEV) is a promising transportation technology with regard to the objective of reducing the exhaust gazes emission, which maintains a relatively high autonomy compared to EV [20].

O. Gusikhin and K. Madani (Eds.): ICINCO 2017, LNEE 495, pp. 234–251, 2019.
https://doi.org/10.1007/978-3-030-11292-9_12

Concerning the energy optimization in the HEV and EV, the researchers mainly deal with two kind of problems: (1) energy power management for a given velocity profile [6,16,26]; (2) velocity profile optimization for EV or conventional vehicles [9,22,32]. The application of the optimal control theory to power management on HEV has been the most popular approach, which includes linear programming, optimal control and especially Dynamic Programming (DP) [2,6,14,21,23,26,28]. These techniques have been widely studied and applied to a broad range of vehicles.

The authors in [23] propose an approach for determining the battery State Of the Charge (SOC_{bat})-dependent equivalent cost factor in HEV supervisory control problems using DP. Song *et al.* [28] use the DP approach to deal with the global optimization problem for deriving the best configuration for the drivetrain components sizes and energy split strategies of a hybrid energy storage system, including a battery and a supercapacitor, for an electric city bus. Authors in [38] proposed a DP-rule based (DP-RB) algorithm to solve the global energy optimization problem in a real time controller of a plug-in hybrid electric bus (PHEB). A control grid (a set of deterministic rules) is built for a typical city route according to the station locations and discrete SOC_{bat} levels. An offline DP with historical running information of the driving cycle is used to deduce optimal control parameters of RB on all points of the control grid.

Other authors as Tokekar *et al.* [32] studied the problem of the velocity profiles search for a car-like robots in order to minimize the energy consumed while traveling along a given path, whereas Dib *et al.* [9] tackle an energy management problem for an electric vehicle compliant with online requirements for "eco-driving application". The main difference between two last papers cited above is that the robot is fully autonomous, and the electric vehicle is controlled by a driver, but the driver receives the velocity profile proposed by an eco-driving system. Authors in [22] propose an optimization of the speed trajectory to minimize the fuel consumption and communicate it to the driver. In their approach the driver sends the information of the intended travel destination to the server. The server generates a route, collects the associated traffic and geographical information, and solves the optimization problem by a spatial domain DP algorithm that utilizes accurate vehicle and fuel consumption models to determine the optimal speed trajectory along the route. Kim *et al.* [15] use model predictive control for the velocity and power split optimization in HEVs. A given velocity profile is optimized by setting the constraints on the velocity and the acceleration of the vehicle. This allows to smooth the current velocity profile without generating a new one. The authors in [34] proposed a method that solves the velocity optimization problem for HEVs, based upon information from Global Navigation Satellite-based Systems, assuming that the velocity trajectory has a predefined shape. Although this method is used for HEVs, the authors do not deal with the energy management optimization aspect. A sub-optimal strategy, based on Deterministic Dynamic Programming (DDP) approach, for online energy optimization of a HEV is proposed in [1]. It uses mainly an appropriate speed profile and power-split database, obtained offline with DDP, in order to cope with different traffic situations, and this is carried out by using a multi-dimensional interpolation method. This approach permits to have simultaneous speed profile optimization and optimal power split strategy.

In effective transportation systems, a totally deterministic model is unlikely to include various uncertainties, associated with sensor's measurement errors, power demand with stochastic nature. To overcome this drawback, some researchers make use of Stochastic Dynamic Programming (SDP) [10,13]. The basic idea behind SDP is the fact that the driver behavior can be modeled and predicted by a Stationary Markov Chain [24]. Plotting the data in the speed-power plane reveals that the driver behavior can be predicted to some extent [10]. Authors in [13] carried out a study to investigate what can be achieved if all the available on-board information is used optimally to assess the potential of predictive control for HEV powertrains. The DDP is used for dimensioning of the energy storage elements and the SDP is used to find a causal operating strategy as in [25]. A time-invariant state feedback based control law, derived from SDP, is proposed in [19] as a power management strategy.

According to [18], a battery SOC_{bat} gradual decrease to a lower threshold leads to better fuel economy, compared with Charge-Depleting and Charge-Sustaining (CDCS) strategy. Authors in [12], as well as in [27,29], deal with the problem of prediction of the battery SOC_{bat}. These papers use an offline global optimal control to generate the desired SOC_{bat} trajectory, later these values are used as an input in Model Predictive Control (MPC). It is proved that prediction of the future trajectories, based upon either past or predicted vehicle velocity and road grade trajectories, could help in obtaining a solution close to the optimal one [35]. Tulpule *et al.* [33] assume the battery SOC_{bat} is linearly decreasing with the distance traveled. A Kalman Filter based Estimator is proposed in [1] to predict a reference SOC_{bat}.

Unlike the above cited references, the present paper proposes an optimization technique based on SDP permitting to have simultaneous speed profile and powersplit optimization strategy for a series-parallel hybrid bus. As driver's power demand is influenced by various driving conditions, it is presented as a random Markov process. An online sub-optimal speed profile and related powersplit generation is developed to deal online with the current road profile and driver power demand. This is carried out using an Optimal Profile Database based on SDP (OPD-SDP), where different transition probabilities and rewards are collected and utilized depending on reference battery SOC_{bat} curve, generated in advance by training an Artificial Neural Network based module for several standard driving cycles. The reference SOC_{bat} curve constraint guarantees a smooth battery discharge so that also at the end of the operational cycle (in the end of a course of a day) the SOC_{bat} do not fall below its permitted minimum level.

The rest of the paper is organized as follows. In Sect. 2, the studied bus powertrain and its dynamical model are presented. Section 3 presents the stochastic modeling of the process, cost function definition and reference SOC_{bat} generation method. Section 4 describes SDP based control strategy. In Sect. 5, several simulation results are presented showing the efficiency of the proposed velocity profile optimization and stochastic energy management strategies. Finally, conclusions and some prospects are given in the last section.

2 Modeling of the Hybrid Bus

The aim of this section is to illustrate the architecture and the mathematical model of the studied system, i.e., BUSINOVA hybrid bus, developed by SAFRA company[1]. This bus is composed of an electric motor, a hydraulic motor, an internal combustion engine and battery as the propulsion powertrain system of the vehicle.

2.1 Hybrid Bus Powertrain Architecture

The model of the studied hybrid bus is based on a series-parallel power-split hybrid architecture [4]. A simple block diagram of the power flows in the bus is shown in Fig. 1. The electric (EM) and hydraulic (HM) motors are both directly connected to the transmission and can ensure simultaneously or independently the traction of the bus. On the other hand, the internal combustion engine (ICE) is coupled to a hydraulic pump (HP) for driving the HM, and therefore allowing the ICE load shifting.

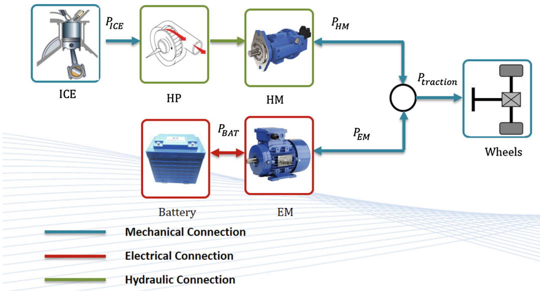

Fig. 1. Block diagram of the powertrain power flows. (ICE: internal combustion engine, HP: hydraulic pump, HM: Hydraulic motor, EM: electric motor) [1].

The rotational speeds of the HM and the EM are imposed by the wheels speed in proportion to the reduction ratios of HM and EM respectively. Moreover, the rotational speed ω_{HM} and the torque T_{HM} are expressed as follows:

$$\begin{cases} \omega_{HM}\left(T_{ICE}, D_{HM}\right) = \frac{D_{HP}.\eta_{v_{HM}}.\omega_{ICE}}{D_{HM}.\eta_{v_{HP}}} \\ T_{HM}\left(T_{ICE}, D_{HM}\right) = \frac{D_{HM}.\eta_{m_{HM}}.T_{ICE}}{D_{HP}.\eta_{m_{HP}}} \end{cases} \tag{1}$$

where ω_{ICE}, T_{ICE} are respectively rotational speed and torque of the ICE, and D_{HM}, D_{HP}, $\eta_{m_{HM}}$, $\eta_{m_{HP}}$, $\eta_{v_{HM}}$, $\eta_{v_{HP}}$ are respectively displacement, mechanical efficiency and volumetric efficiency of the HM and the HP.

[1] http://www.businova.com

The BUSINOVA can operate according to the modes described below:

1. the propulsion is fully supplied by the electric motor (mode I),
2. the bus is actuated by the hydraulic motor via the ICE (mode II),
3. the mode III implies the hybrid operation of the EM and the HM via ICE,
4. the regenerative braking (mode IV) - the part of the kinetic energy during braking phase is recuperated to charge the electric battery.

2.2 Dynamical Model

This part is dedicated to the dynamical equations describing the bus. The purpose of the dynamical model is to have a realistic global behavior of the bus in order to validate the proposed energy management techniques. To describe it in a generic manner, assume that the bus is moving up the slope of θ degree (cf. Fig. 2). The origin of the coordinates is situated in the Center of Mass (CoM). It is supposed that CoM of the bus is in its geometric center. The dynamical equation of the bus is given as follows:

$$F_{tr} + F_{rr} + F_{ad} + F_g + F_{brake} = (M + M_{eq})a \tag{2}$$

where $\boldsymbol{F}_{tr}$ traction force, $\boldsymbol{F}_{rr}$ rolling resistance, $\boldsymbol{F}_{ad}$ aerodynamic force, $\boldsymbol{F}_g$ gravity force, $\boldsymbol{F}_{brake}$ mechanical brake force, M bus weight, M_{eq} equivalent mass of rotating parts, $\boldsymbol{a}$ bus acceleration.

To produce the bus acceleration, it is necessary to take into account the moments of inertia of the rotating components (e.g., rotor of the EM, crankshaft of the ICE, driving axle, etc.). It is done by introducing the equivalent mass M_{eq} of the rotating components:

$$M_{eq} = \frac{i_g \eta_{pt} J_{rot}}{r^2} \tag{3}$$

where i_g gear ratio, η_{pt} powertrain efficiency, J_{rot} total inertia of the rotating components in the transmission, and r the wheel radius [7].

The traction force F_{tr} is linked to the torque produced by the powertrain T_{pt} via gear ratio i_g, powertrain efficiency η_{pt}. Expanding the dynamical equation (2), the following relation is obtained:

$$a = \frac{dv}{dt} = \frac{1}{M + M_{eq}} H \tag{4}$$

with

$$H = \frac{i_g \eta_{pt} T_{pt}}{r} - \mu_{rr} F_N sign(v) - \frac{1}{2}\rho A C_d (v + v_{wind})^2 - Mg\sin(\theta) - \frac{T_{brake}}{r} \tag{5}$$

where:

- T_{pt}: output powertrain torque from the gearbox,
- μ_{rr}: rolling resistance coefficient, $F_N = Mgcos(\theta)$ normal force, g gravity acceleration, θ slope angle, v bus speed,
- ρ: the air density, A the frontal area of the bus, C_d drag coefficient, v_{wind} wind speed,
- T_{brake}: the brake torque provided by the bus mechanical brake system.

3 Overall Multi-criteria Optimization Formulation

The objective of an optimal control problem is to find the optimal bus velocity profile and energy split between the actuators for a given trajectory D. Section 3.1 describes a proposed global control architecture. Section 3.2 is dedicated to the stochastic modeling of driver's behavior. Section 3.3 describes an approach for reference SOC_{bat} curve generation, and Sect. 3.4 defines a proposed multi-objective minimization criteria formulation.

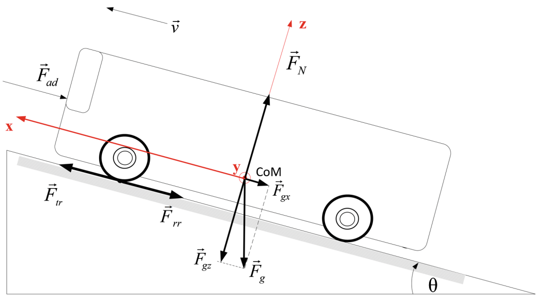

Fig. 2. Forces acting on the bus. [1].

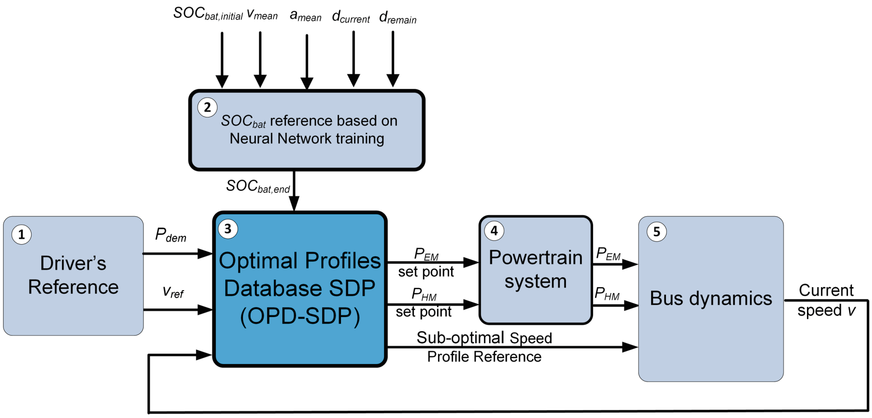

Fig. 3. Proposed global control architecture for online stochastic sub-optimal speed profiles and its powersplit generation.

3.1 Global Control Architecture

A global stochastic control architecture is proposed in Fig. 3.

It consists of 5 blocks:

- **Block** ① calculates the demanded power P_{dem} of the vehicle for a reference driver's speed profile.
- **Block** ② calculates the reference SOC_{bat} curve (cf. Sect. 3.3).
- **Block** ③ stands for Optimal Profiles Database based on Stochastic Dynamic Programming. This block contains Transition Probability (cf. Sect. 3.2) and Reward Matrices, based on the defined cost function (cf. Sect. 3.4). An optimal speed profile and powersplit are obtained using Stochastic Dynamic Programming approach which is detailed in Sect. 4.
- **Block** ④ and **Block** ⑤ contain dynamic model of the bus and the powertrain architecture dynamics.

3.2 Stochastic Modeling of Driver

The driver's throttle and brake pedal commands are interpreted as a power demand to be satisfied by the powertrain. The driver's behavior can be considered as a stochastic process [3,37], therefore the driver model is represented as a Markov Decision Process in this paper. The system is described by the state variable $x = \{v, P_{dem}, SOC_{bat}\}$, where v bus speed, P_{dem} total power demand, and SOC_{bat} battery state of charge. The power distribution between two energy storages and the speed profile optimization is controlled by the control signal $u = \{a, P_{EM}\}$, where a is acceleration and P_{EM} electric motor power. Based on collected and simulated data of driving cycles, the transition probability of the state variable are calculated and modeled as a homogeneous Markov Chain.

The state and control spaces are discretized so that they take a finite number of values. The state variables are discretized as follows:

$$P_{dem} \in \{P_{dem}^1,\ P_{dem}^2,\ \ldots,\ P_{dem}^{N_p}\} \tag{6}$$

$$v \in \{v^1,\ v^2,\ \ldots,\ v^{N_v}\} \tag{7}$$

$$SOC_{bat} \in \{SOC_{bat}^1,\ SOC_{bat}^2,\ \ldots,\ SOC_{bat}^{N_E}\} \tag{8}$$

So the total state space dimension is:

$$\{x^i,\ i = 1,\ 2,\ \ldots,\ N_p N_v N_E\} \tag{9}$$

The control variables u are also discretized:

$$a \in \{a^1,\ a^2,\ \ldots,\ a^{N_a}\} \tag{10}$$

$$P_{EM} \in \{P_{EM}^1,\ P_{EM}^2,\ \ldots,\ P_{EM}^{N_{EM}}\} \tag{11}$$

Since the P_{dem} is assumed to be a Markovian state, the power dynamics is defined as follows:

$$P_{dem,k+1} = w_k \tag{12}$$

where the probability distribution of w_k is assumed to be:

$$Pr\{w = P^j_{dem} \mid P_{dem} = P^i_{dem},\ v = v^l\} = p_{il,j} \tag{13}$$

$i, j = 1,\ 2,\ \ldots,\ N_p$, $l = 1,\ 2,\ \ldots,\ N_v$, where $\sum_{j=1}^{N_p} p_{il,j} = 1$, and $p_{il,j}$ represents the transition probability of the system in a state P^j_{dem} at instant $k+1$ with regards to the system in state P^j_{dem} and v^l at instant k.

We used standard driving cycles in this study to determine the transition probabilities as follows: various standard driving cycles were selected to represent different urban driving cycles. From the reference speed profile, P_{dem} has been calculated using the vehicle model (cf. Sect. 2). Using nearest-neighbor quantization, the sequence of observations $(P_{dem},\ v)$ was mapped into a sequence of quantized states $(P^i_{dem}, v = v^l)$. The transition probability has been estimated by the maximum likelihood estimator [36], which counts the observation data as:

$$\hat{p}_{il,j} = \frac{m_{il,j}}{m_{il}} \tag{14}$$

with $m_{il} \neq 0$, where $m_{il,j}$ is the number of occurrences of the transition from P^i_{dem} to P^j_{dem}, when the speed was at v^l, $m_{il} = \sum_{j=1}^{n} m_{il,j}$ is the total number of times that P^i_{dem} has occurred at speed v^l.

3.3 Reference Battery SOC

The BUSINOVA bus is a plug-in hybrid electric vehicle and its standard functioning time is 8 h a day (so called "course of a day"). Figure 4 illustrates the spatial bounds of a bus running cycle. The bus travels from its starting location to another route terminus, stopping at bus stations (BS) along the route to allow

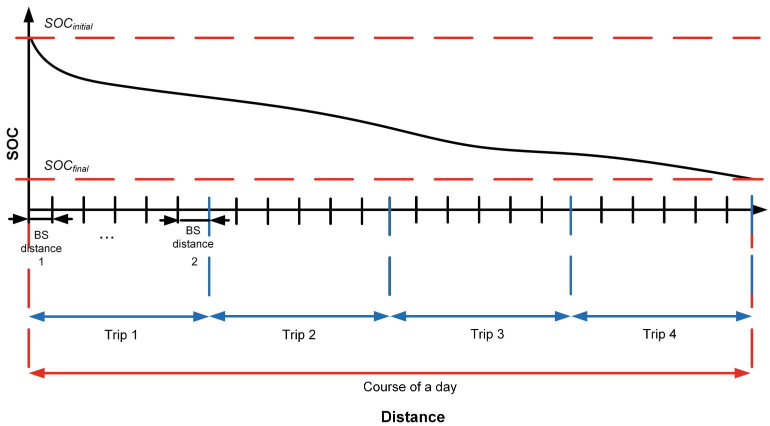

Fig. 4. Spatial bounds of a bus running cycle.

passengers to board and to alight. This movement is called a trip. By the end of a day, the bus reaches its SOC_{min} value and can be recharged during all the night long to ensure the service the next day.

In this work, the principle idea is to consider that a better usage of the electric energy is such that it is available until the end of the day (during 8 h operational cycle), and this is considered as an optimal functioning of the bus. The working hypothesis behind this assumption is to use the maximum amount of energy that can be consumed from the battery in one day driving so that the battery energy is spread as uniformly as possible in one working day. This implies the smooth battery discharging rate (C-rate), avoidance of the high or low SOC_{bat} and excessive depth of charge, which lead to a high rate of battery capacity loss [5,8,30]. As Li-ion batteries represent a big part of a vehicle cost, the clear interest is to prolongate the battery life. For that purpose a SOC_{bat} Estimator based on Kalman Filter has been proposed in [1].

An Artificial Neural Network (ANN) module was designed to learn the mechanisms of optimal SOC_{bat} curves according to different trip information [31]. After getting the average speed and length of route segments when the trip starts, the ANN module is used to generate a reasonable and relatively precise SOC_{bat} reference curve. However, frequent acceleration phases strongly impact the energy consumption in the vehicle [17]. Thus, in this work, an average absolute acceleration value has been added as an input to the ANN module, in order to take into account the dynamics of the speed variation on a given route segment. The block ② in Fig. 3 corresponds to the SOC_{bat} curve learning module based on ANN which has as inputs: $SOC_{bat,initial}$ initial battery SOC_{bat}, v_{mean} average speed, a_{mean} average absolute acceleration value, $d_{current}$ traveled distance, and d_{remain} remained distance. This relation is illustrated in Fig. 5. The road is divided into segments D_n. We can see that the more important is the average speed/acceleration (on a given route segment), the more energy is consumed in the corresponding phase.

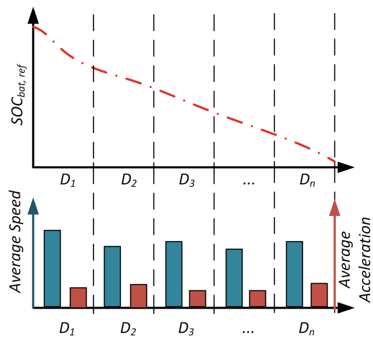

Fig. 5. Relation among average speed/acceleration and SOC_{bat} reference.

3.4 Multi-objective Optimization Problem Formulation

As any problem of optimization, it is important to define optimization criteria. In our case, the criteria is defined in order to minimize the energy consumed by the HEV while optimizing both: HEV speed profile and the powersplit during the trip D. In previous work [1] a minimizing cost function consisted in a compromise between electric motor and engine consumed power. The weighted aggregation method has been used to find a set of sub-optimal trade-off solutions of the cost function. In this work, a driver's power demand is considered as a random MDP (cf. sub-sect. 3.2). Thus, the following cost function ζ is proposed and formulated as a Multi-Objective (MO) optimization problem:

$$\zeta = f_1 + f_2 \tag{15}$$

where f_1 is a criterion responsible for speed optimization and f_2 for powersplit optimization. The given criteria are defined as follows:

$$f_1 = (\frac{v_{ref} - v}{dt} - a)^2 \tag{16}$$

$$f_2 = (SOC(D_n) - SOC_{current})^2 + (P_{dem} - P_{EM} - P_{HM})^2 \tag{17}$$

where v_{ref} is a reference driver speed, v bus current speed, dt time interval, a acceleration control input, $SOC(D_n)$ reference SOC_{bat} value for a route segment D_n, $SOC_{current}$ current SOC_{bat} value, P_{dem}, P_{EM} and P_{HM} demanded power, electric and hydraulic motors supplied power, respectively.

The goal is to find a set of optimal solutions that minimize f_1 and f_2 among all the feasible solutions, i.e., generate a Pareto front (cf. Fig. 6). To this end, ϵ-constraint method was chosen. It is a MO optimization technique, proposed by Haimes *et al.* [11], for generating Pareto optimal solutions. It makes use of a single-objective optimizer which handles constraints, to generate one point of the Pareto front at a time. For transforming the MO problem into several single-objective problems with constraints the authors use the following procedure (assuming minimization for all the objective functions):

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_l(x) \\ \text{subject to} \quad & f_j(x) \leq \epsilon_j, \; i = 1, \ldots, m, j \neq l \\ & x \in S \end{aligned}$$

where $l \in \{1, 2, ..., m\}$ and S is the feasible region, which can be defined by any equality and/or inequality constraint. The vector of upper bounds, $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_m)$, defines the maximum value that each objective can have. In order to obtain a subset of the Pareto optimal set (or even the entire set, in case if this set is finite), one must vary the vector of upper bounds along the Pareto front for each objective, and make a new optimization process for each new vector.

Based on the defined cost function, the Reward Matrices are calculated and stored in an Optimal Profile Database based on Stochastic Dynamic Programming (OPD-SDP) in order to be used to find optimal policy for infinite horizon problem.

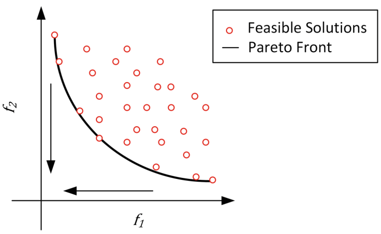

Fig. 6. Pareto Front of the minimization functions.

4 Sub-optimal Energy Management Strategy Using Database Based on SDP

After a Markov model has been built, Transition Probability and Reward Matrices are stored in the OPD-SDP, the Stochastic Dynamic Programming is used to find the optimal control, minimizing the expected cost function (15), for each vehicle state. The control signal $u = \{a, P_{EM}\}$. An infinite horizon problem is formulated on the homogeneous Markov chain. The optimization finds an optimal policy, $u = \pi(x)$ that minimizes the expected cost function (15) over an infinite horizon [24].

$$J_\lambda^\pi(x_0) = \lim_{N \to \infty} E\{\sum_{k=0}^{N-1} \lambda^k \zeta(x_k, \pi(x_k))\} \tag{18}$$

where ζ is the cost for one time step, λ is the discount factor, x_k is the dynamic state vector at the k^{th} time point and $E\{\dots\}$ denotes the expectation with respect to the considered prediction model defined by the time invariant Markov chain. The discount factor $\lambda < 1$ assures convergence of the infinite sum. The optimal policy $u = \pi(x)$ is found by using a modified policy iteration algorithm (cf. Fig. 7). The modified policy iteration consists of the following steps:

1. Initial guess: Set $i = 1$. Provide an initial guess for $\pi_i(x)$ and the discounted infinite horizon future cost $J_\lambda^{\pi_{i-1}}(x)$.
2. Policy evaluation: Calculate the cost $J_\lambda^{\pi_i}(x)$ of the policy $\pi_i(x)$ by iterating

$$J_\lambda^{\pi i}(x_k) = \zeta(x_k, u) + E\{\lambda J_\lambda^{\pi i}(x_{k+1})\}$$

backwards N times starting from $J_\lambda^{\pi_i}(x_N) = J_\lambda^{\pi_{i-1}}(x)$ and ending with the truncated cost $J_\lambda^{\pi_i}(x) = J_\lambda^{\pi_i}(x_0)$. If $\mid J_\pi^{i+1}(x) - J_\pi^i(x) \mid \leq \xi$ when iterating backwards terminate the iterations with the answer $J_\lambda^{\pi_i}(x) = J_\lambda^{\pi_i}(x_k)$.
3. Policy improvement: Improve the policy by taking one value iteration step:

$$\pi_{i+1}(x_0) = \arg\min_u \left[\zeta(x_0, u) + E\{\lambda J_\lambda^{\pi_i}(x_1)\right].$$

Check for convergence: if $\mid J_\pi^{i+1}(x_0) - J_\pi^i(x_k) \mid \leq \xi$, then terminate the algorithm with the answer $\pi(x) = \pi_{i+1}(x)$. If $\pi_{i+1}(x)$ has not converged, increase the index: i = i+1, and go to step 2.

The flowchart of the Modified Policy Iteration Algorithm for SDP is presented in Fig. 7. The given algorithm efficiency is tested and validated for real-time application. The simulation results are presented in Sect. 5.

The calculation of SDP consumes relatively big amount of memory, so a reasonable resolution of state variables number is very important in terms of computation effort. With trials and errors method, the following discretization sets for state and action variables have been applied:

$$P_{dem} = \{-200 : 20 : 160\}\text{kW} \tag{19}$$

$$v = \{0,\ 5,\ 10,\ 20\}\text{m/s} \tag{20}$$

$$SOC_{bat} = \{0.2 : 0.1 : 0.9\} \tag{21}$$

$$a = \{-3, -1, 0, 0.5, 1, 1.25, 1.5\}m/s^2 \tag{22}$$

$$P_{EM} = \{-10 : 5 : 60\}\text{kW} \tag{23}$$

Based on the defined sets, the Transition Probability and Reward Matrices are calculated.

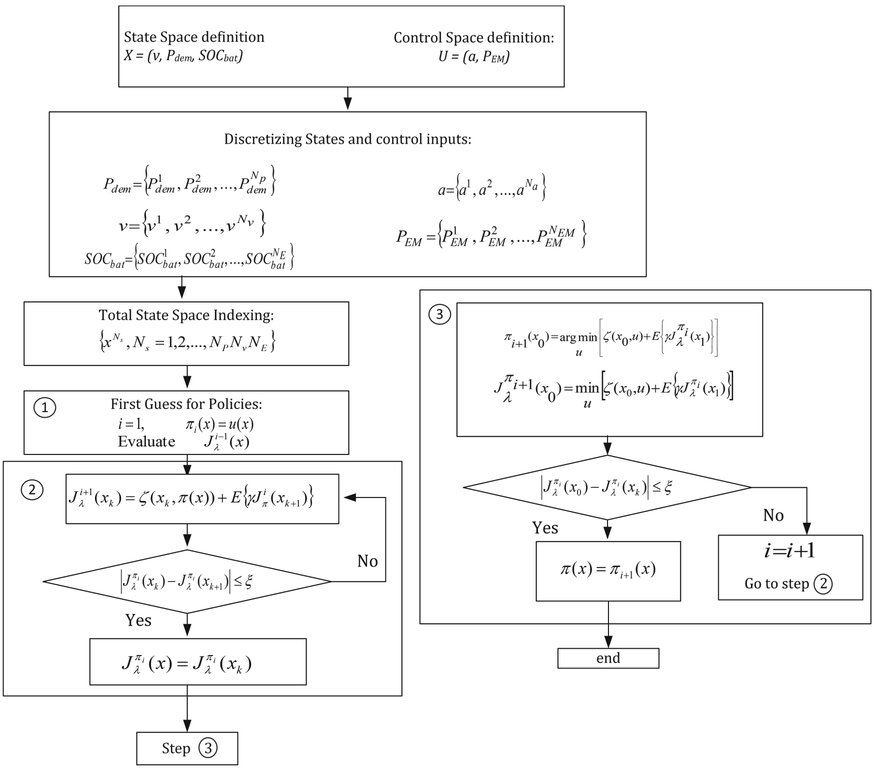

Fig. 7. Modified policy iteration algorithm flowchart.

4.1 Transition Probability Matrices Generation

As it was aforementioned in Sect. 3.2, the Transition Probability Matrices (TPM) are calculated using maximum likelihood estimation (cf. Eq. (14)). In order to have a richer database for TPM estimation, data from the following standard driving cycles (SDC) have been collected: ECER15 (also known as UDC - Urban Driving Cycle), EUDC (European Urban Driving Cycle), ArtUrban (Urban Artemis), and NEDC (New European Driving Cycle). These SDC represent common driving conditions in an urban environment, emulating different driver's behavior (abrupt and smooth acceleration/deceleration, cruise speed phases, etc).

The variation of bus weight significantly influences the vehicle dynamics (thus its consumed energy), as it can vary up to several tons. The tested data has been enlarged by simulating each of the SDC for three different masses, corresponding to an empty, half-full, and full bus. Therefore, the bus dynamics will be tested for much more states, as for the same speed profile the power demand will be different. It is considered in these hypotheses that a full bus corresponds to 70 passengers with an average weight of 60 kg.

After obtaining the data from several SDC for different masses, it was concluded that no all states are explored and some of them are barely feasible. That is why it makes sense to eliminate those states and, thus, reduce the state dimension.

5 Simulation Results

In order to evaluate the efficiency of the proposed approach, a validation scenario has been performed for several driving cycles. The given stochastic strategy has been tested and compared to a strategy based on DDP based database as given in [1]. The validation scenario consists of a test with an assumption that bus mass varies during the trip. The bus is a transportation mean, it is considered that the mass changes only during stops in order to take or to drop off the passengers.

Table 1 summarizes the obtained results for several SDC. "+" corresponds to a positive improvement, "-" means that DDP approach demonstrated a better performance. The consumed energy E_{cons} given by the column **Energy [kWh]** in Table 1 is calculated as follows:

$$E_{cons} = \int_0^{t_f} P_{EM} dt + \int_0^{t_f} Q_{LHV} \dot{m}_f dt \tag{24}$$

where P_{EM} is the electric motor consumed power, $\dot{m}_f$ fuel consumption rate, and $Q_{LHV} = 43\,\text{MJ/kg}$ lower heat value for diesel.

In order to validate the proposed algorithm, in terms of generalization aspect, it has been additionally tested on an ArtRoad SDC, which has not been used before to estimate the TPM. To illustrate the performance, the simulation results for ArtRoad and ArtUrban SDC are presented in Figs. 8 and 9, respectively.

Figure 8 shows the simulation results for ArtRoad SDC for the case when it is assumed that the bus weight is variable. We can see that the stochastic algorithm shows results close to the deterministic approach. Global energy consumed during the trip while using DDP approach is 4.95% higher comparing to SDP method. It outperforms the DDP approach due to smoother speed profile and energy management algorithm.

Figure 9 shows the simulation results for ArtUrban SDC. This is one of the profiles that has been used to train the TPM. However, in this case the bus weight is variable as well. In this more realistic scenario, the SDP approach still outperforms the DDP method by 5.68%.

Both ArtRoad and ArtUrban SDC are characterized by frequent acceleration and deceleration phases. In this case, the SDP based approach demonstrates better results. For the SDC with mainly constant speed phases (e.g., EUDC and NEDC) the improvement compared to DDP approach is more modest or even negative (cf. Table 1). Therefore, the expediency of the use of the SDP app-

Table 1. Comparison of the SDP w.r.t. DDP based strategies. Variable weight.

Driving Cycle	ΔSOC [%]			Fuel [l]			Energy [kWh]		Energy Diff. [%]
	DDP	SDP	Diff. [%]	DDP	SDP	Diff. [%]	DDP	SDP	
ArtRoad	8.15	7.91	+2.91	0.48	0.47	+2.0	8.22	7.81	+4.98
EUDC	4.51	4.45	+1.33	0.11	0.12	−9.1	2.71	2.64	+2.58
ArtUrban	2.15	1.87	+13.3	0.60	0.59	+1.67	7.04	6.64	+5.68
NEDC	4.60	4.66	−1.3	0.15	0.17	−13.3	3.65	3.85	−5.47
Average	–	–	+4.06	–	–	−4.68	–	–	+1.94

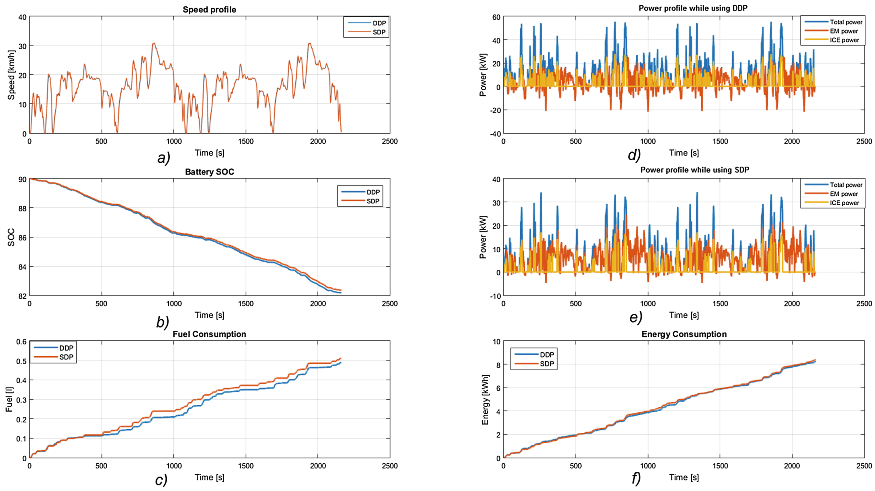

Fig. 8. ArtRoad SDC. Constant weight. *(a)* Speed profile, *(b)* Battery SOC curve, *(c)* Fuel consumption curve, *(d)* Powersplit for DDP approach, *(e)* Powersplit for the proposed SDP approach, *(f)* Total energy consumption.

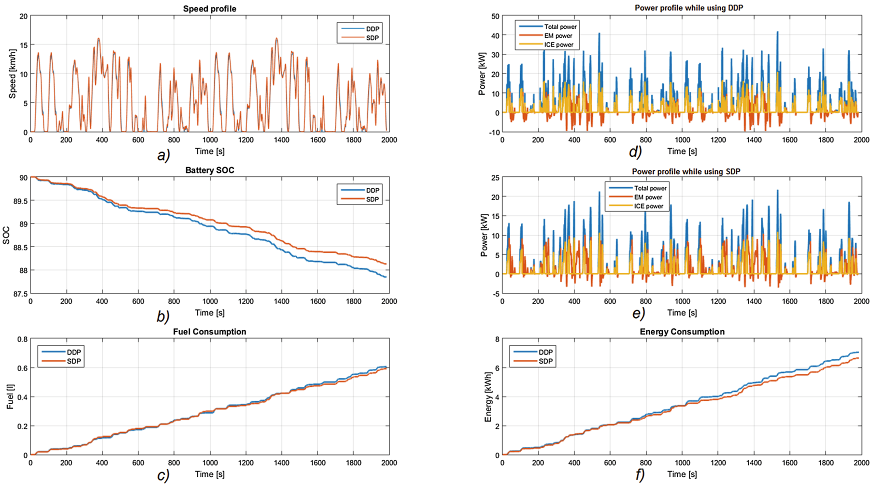

Fig. 9. ArtUrban SDC. Variable weight. *(a)* Speed profile, *(b)* Battery SOC curve, *(c)* Fuel consumption curve, *(d)* Powersplit for DDP approach, *(e)* Powersplit for the proposed SDP approach, *(f)* Total energy consumption.

roach is more advantageous in case if a vehicle is subject to some uncertainties, e.g., aggressive driving style and/or traffic conditions (traffic lights, traffic jams, pedestrian crossing, etc.), which lead to frequent acceleration and deceleration phases.

6 Conclusion and Prospects

In this paper, a Stochastic Dynamic Programming technique is used to simultaneously generate an optimal speed profile and related powersplit strategy for HEV, in order to ensure energy optimization in the presence of uncertainties (due to different traffic conditions and/or a driver's behavior). To this end, a driver's power demand is modeled as a Markov Chain. The formulated energy optimization problem, being intrinsically multi-objective problem, has been transformed into several single-objective ones with constraints using an ϵ-constraint method to determine a set of optimal solutions that represent the Pareto Front.

Simulation data of several standard urban driving cycles have been used to train Transition Probability Matrices (TPM) using the maximum likelihood estimation technique. A trained reference SOC_{bat} curve has been utilized as a constraint to estimate Reward Matrices based on the defined cost function. For a real-time application purpose, the obtained TPM and Reward Matrices have been collected into an Optimal Profile Database based on SDP (OPD-SDP). The OPD-SDP has been then used to calculate sub-optimal speed profiles and related powersplit by Modified Policy Iteration Algorithm using an infinite horizon optimization formulation. The results obtained by the SDP were compared to a Deterministic Dynamic Programming, and it has been shown that near optimal results can be obtained in real-time application. Later, the given approach will be implemented in the actual studied bus.

References

1. Abdrakhmanov, R., Adouane, L.: Dynamic programming resolution and database knowledge for online predictive energy management of hybrid vehicles. In: International Conference on Informatics in Control, Automation and Robotics (ICINCO), Madrid-Spain (2017)
2. Abdrakhmanov, R., Adouane, L.: Efficient ACC with stop and go maneuvers for hybrid vehicle with online sub-optimal energy management. In: RoMoCo'17, 11th International Workshop on Robot Motion and Control, Wasowo-Poland, 3–5 July 2017
3. Angkititrakul, P., Miyajima, C., Takeda, K.: An improved driver-behavior model with combined individual and general driving characteristics. In: 2012 IEEE Intelligent Vehicles Symposium (IV), pp. 426–431. IEEE (2012)
4. Bayindir, K.Ç., Gözüküçük, M.A., Teke, A.: A comprehensive overview of hybrid electric vehicle: powertrain configurations, powertrain control techniques and electronic control units. Energy Convers. Manag. **52**(2), 1305–1313 (2011)
5. Broussely, M., et al.: Main aging mechanisms in LI ion batteries. J. Power Sources **146**(1), 90–96 (2005)
6. Chen, B.C., Wu, Y.Y., Tsai, H.C.: Design and analysis of power management strategy for range extended electric vehicle using dynamic programming. Appl. Energy **113**, 1764–1774 (2014)
7. Cheng, Y., Lataire, P., et al.: Research and test platform for hybrid electric vehicle with the super capacitor based energy storage. In: 2007 European Conference on Power Electronics and Applications, pp. 1–10. IEEE (2007)
8. Choi, S.S., Lim, H.S.: Factors that affect cycle-life and possible degradation mechanisms of a LI-ion cell based on $LICoO_2$. J. Power Sources **111**(1), 130–136 (2002)
9. Dib, W., Chasse, A., Moulin, P., Sciarretta, A., Corde, G.: Optimal energy management for an electric vehicle in eco-driving applications. Control. Eng. Pract. **29**, 299–307 (2014)
10. Elbert, P., Widmer, M., Gisler, H.J., Onder, C.: Stochastic dynamic programming for the energy management of a serial hybrid electric bus. Int. J. Veh. Des. **69**(1–4), 88–112 (2015)
11. Haimes, Y.Y.: On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Trans. Syst. Man Cybern. **1**(3), 296–297 (1971)
12. Heppeler, G., Sonntag, M., Wohlhaupter, U., Sawodny, O.: Predictive planning of optimal velocity and state of charge trajectories for hybrid electric vehicles. Control. Eng. Pract. **61** (2016)
13. Johannesson, L., Asbogard, M., Egardt, B.: Assessing the potential of predictive control for hybrid vehicle powertrains using stochastic dynamic programming. IEEE Trans. Intell. Transp. Syst. **8**(1), 71–83 (2007)
14. Kamal, E., Adouane, L., Abdrakhmanov, R., Ouddah, N.: Hierarchical and adaptive neuro-fuzzy control for intelligent energy management in hybrid electric vehicles. In: IFAC World Congress. Toulouse-France (2017)

15. Kim, T.S., Manzie, C., Sharma, R.: Model predictive control of velocity and torque split in a parallel hybrid vehicle. In: 2009 IEEE International Conference on Systems, Man and Cybernetics, SMC 2009, pp. 2014–2019. IEEE (2009)
16. Kitayama, S., Saikyo, M., Nishio, Y., Tsutsumi, K.: Torque control strategy and optimization for fuel consumption and emission reduction in parallel hybrid electric vehicles. Struct. Multidiscip. Optim. **52**(3), 595–611 (2015)
17. Larsson, H., Ericsson, E.: The effects of an acceleration advisory tool in vehicles for reduced fuel consumption and emissions. Transp. Res. Part D Transp. Environ. **14**(2), 141–146 (2009)
18. Li, L., Yang, C., Zhang, Y., Zhang, L., Song, J.: Correctional DP-based energy management strategy of plug-in hybrid electric bus for city-bus route. IEEE Trans. Veh. Technol. **64**(7), 2792–2803 (2015)
19. Lin, C.C., Peng, H., Grizzle, J.: A stochastic control strategy for hybrid electric vehicles. In: Proceedings of the 2004 American Control Conference, 2004, vol. 5, pp. 4710–4715. IEEE (2004)
20. Murphey, Y.L.: Intelligent vehicle power management: an overview. Computational Intelligence in Automotive Applications, pp. 169–190. Springer, Berlin (2008)
21. Ouddah, N., Adouane, L., Abdrakhmanov, R., Kamal, E.: Optimal energy management strategy of plug-in hybrid electric bus in urban conditions. In: International Conference on Informatics in Control, Automation and Robotics (ICINCO), Madrid-Spain (2017)
22. Ozatay, E., et al.: Cloud-based velocity profile optimization for everyday driving: a dynamic-programming-based solution. IEEE Trans. Intell. Transp. Syst. **15**(6), 2491–2505 (2014)
23. Pei, D., Leamy, M.J.: Dynamic programming-informed equivalent cost minimization control strategies for hybrid-electric vehicles. J. Dyn. Syst. Meas. Control. **135**(5), 051,013 (2013)
24. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New Jersey (2014)
25. Romaus, C., Gathmann, K., Böcker, J.: Optimal energy management for a hybrid energy storage system for electric vehicles based on stochastic dynamic programming. In: 2010 IEEE Vehicle Power and Propulsion Conference (VPPC), pp. 1–6. IEEE (2010)
26. Rousseau, G.: Véhicule hybride et commande optimale. Ph.D. thesis, École Nationale Supérieure des Mines de Paris (2008)
27. Shen, D., Bensch, V., Miiller, S.: Model predictive energy management for a range extender hybrid vehicle using map information. IFAC-PapersOnLine **48**(15), 263–270 (2015)
28. Song, Z., Hofmann, H., Li, J., Han, X., Ouyang, M.: Optimization for a hybrid energy storage system in electric vehicles using dynamic programing approach. Appl. Energy **139**, 151–162 (2015)
29. Sun, C., Sun, F., Hu, X., Hedrick, J.K., Moura, S.: Integrating traffic velocity data into predictive energy management of plug-in hybrid electric vehicles. In: 2015 American Control Conference (ACC), pp. 3267–3272. IEEE (2015)
30. Tang, L., Rizzoni, G., Onori, S.: Energy management strategy for hevs including battery life optimization. IEEE Trans. Transp. Electrification **1**(3), 211–222 (2015)
31. Tian, H., Wang, X., Lu, Z., Huang, Y., Tian, G.: Adaptive fuzzy logic energy management strategy based on reasonable SOC reference curve for online control of plug-in hybrid electric city bus. IEEE Trans. Intell. Transp. Syst. (2017)
32. Tokekar, P., Karnad, N., Isler, V.: Energy-optimal trajectory planning for car-like robots. Auton. Robot. **37**(3), 279–300 (2014)

33. Tulpule, P., Marano, V., Rizzoni, G.: Effects of different phev control strategies on vehicle performance. In: 2009 American Control Conference, ACC'09, pp. 3950–3955. IEEE (2009)
34. Van Keulen, T., de Jager, B., Foster, D., Steinbuch, M.: Velocity trajectory optimization in hybrid electric trucks. In: Proceedings of the 2010 American Control Conference, pp. 5074–5079. IEEE (2010)
35. Van Keulen, T., de Jager, B., Kessels, J., Steinbuch, M.: Energy management in hybrid electric vehicles: benefit of prediction. IFAC Proc. Vol. **43**(7), 264–269 (2010)
36. White, H.: Maximum likelihood estimation of misspecified models. Econ. J. Econ. Soc. **50**(1), 1–25 (1982)
37. Zeng, X., Wang, J.: A stochastic driver pedal behavior model incorporating road information. IEEE Trans. Hum. Mach. Syst. **47**(5), 614 (2017)
38. Zhang, Y., Jiao, X., Li, L., Yang, C., Zhang, L., Song, J.: A hybrid dynamic programming-rule based algorithm for real-time energy optimization of plug-in hybrid electric bus. Sci. China Technol. Sci. **57**(12), 2542–2550 (2014)