Shortest Path Stochastic Control for Hybrid Electric Vehicles

Ed Tate[1], J.W. Grizzle[2], Huei Peng[3]

Abstract:

When a Hybrid Electric Vehicle (HEV) is certified for emissions and fuel economy, its power management system must be charge sustaining over the drive cycle, meaning that the battery state of charge (SOC) must be at least as high at the end of the test as it was at the beginning of the test. During the test cycle, the power management system is free to vary the battery SOC so as to minimize a weighted combination of fuel consumption and exhaust emissions. This paper argues that shortest path stochastic dynamic programming (SP-SDP) offers a more natural formulation of the optimal control problem associated with the design of the power management system because it allows deviations of battery SOC from a desired setpoint to be penalized only at key-off. This method is illustrated on a parallel hybrid electric truck model that had previously been analyzed using infinite-horizon stochastic dynamic programming with discounted future cost. Both formulations of the optimization problem yield a time-invariant causal state-feedback controller that can be directly implemented on the vehicle. The advantages of the shortest path formulation include that a single tuning parameter is needed to tradeoff fuel economy and emissions versus battery SOC deviation, as compared to two parameters in the discounted, infinite-horizon case, and for the same level of complexity as a discounted future-cost controller, the shortest-path controller demonstrates better fuel and emission minimization while also achieving better SOC control when the vehicle is turned off. Linear Programming is used to solve both stochastic dynamic programs.

[1] Ed Tate. (Corresponding author) General Motors, Hybrid Powertrain Engineering, Mail Code 483-619-406, 1996 Technology Drive, Troy, Michigan, 48007. Phone: (248) 680-4155. ed.d.tate@gm.com
[2] J.W. Grizzle, Dept of Electrical Engineering and Computer Science, University of Michigan, 48109-2122. grizzle@umich.edu
[3] Huei Peng, Dept of Mechanical Engineering, University of Michigan, 48109-2125. hpeng@umich.edu

Introduction.

To achieve greater fuel efficiency and lower emissions in vehicles using internal combustion engines, there have been two basic approaches used in the powertrain. The first approach is to improve the thermodynamic cycle used to convert fuel into power. Technologies such as cylinder deactivation [1], variable cam phasing [2] and timing [3] , continuously variable transmissions [4], automated shift transmissions [5] and exhaust gas recirculation [6] are all used to improve the instantaneous efficiency and decrease emissions from a powertrain. The second approach is to improve on the system performance of the powertrain by storing and releasing energy over a drive cycle, allowing the engine to operate in more favorable regions, which improves the average thermal efficiency of the engine. One technology for achieving this is the Hybrid Electric Vehicle (HEV) which consists of an electric powertrain coupled to a conventional powertrain.

A conventional vehicle powertrain consists of an engine, transmission, differential and controls. The controls translate a request for torque from the driver into inputs for the engine and transmission that produce that desired torque, while achieving multiple objectives. These objectives include fuel economy, emissions, torque deviations, and component protection. An HEV retains all of the control complexity of a conventional powertrain and introduces new complexity. This increase in complexity occurs because of the additional components and the coupling among those components. These additional components increase the number of states to be controlled, the inputs for control, and the outputs available for feedback. Additionally, there are new control objectives. Due to federal and state regulations [7],[8],[9],[10], the controls must manage battery charge in a manner that satisfies certification requirements. In addition to certification requirements, a manufacturer may have other control objectives including battery life management, system thermal management to prevent component overheating, and minimization of the customer's perceptions of 'odd' behavior. While this design process focuses on fuel economy, these other factors can be incorporated.

1.1) Prior Art

To design control laws for HEV's, several approaches have been taken. The dominant design approach in the literature is some form of rule-based design, relying on engineering intuition and logic. For example, the work in [11] determined the optimal control choice by selecting the best action based on the instantaneous cost, where this instantaneous cost includes an approximation of the cost in fuel for electrical energy from the battery. In [12] a similar scheme is employed. As a means to analyze the properties of an optimal control, Dynamic Programming (DP) is used in [13] and [14]. However, these analyses were not used to generate a causal control law. Reference [15] uses the solution to a DP to design a rule-based controller. In that paper, a DP is solved to find a globally optimal sequence of control inputs for the powertrain. The insights gained from this solution were then used to design a realizable control law and implement it. This design approach achieved significant improvements in fuel economy and emissions reduction over heuristically derived control laws when tested in an Eaton prototype HEV [16].

Utilizing assumptions of perfect future knowledge of the vehicle trajectory, Back [17] proposed a causal Model Predictive Control (MPC) control law. Unfortunately, perfect future knowledge is rarely available. In an engineering problem similar to the HEV energy storage issue, Kummert [18] applied Stochastic MPC to a passive solar building. This approach used a prediction (weather forecast) and a confidence in the prediction to determine an optimal schedule for heating and cooling a building. This would suggest that path prediction with a prediction confidence could be used in a scheme similar to [17], though, no work proposing such a technique has been found.

Recently, Lin at al. [20] proposed Stochastic Dynamic Programming (SDP) as a technique to directly design an HEV control law. This design technique yielded better performance than any heuristically derived rule-based controller previously employed. The primary advantage of this approach is that the control law is designed from an optimization criterion, a model of the plant dynamics and a model of the driving pattern statistics. A weakness in this approach is that the optimization criterion discounts future costs and assigns a penalty to SOC deviation from a set point at every instant in time. This optimization contains two tuning parameters, the discount factor and the SOC deviation penalty.

1.2) New Contributions

When the controls for an HEV are developed based on vehicle certification requirements, there is no cumulative cost for deviation from an SOC set point, except when the vehicle is turned off at the end of the test. At that point, there is a penalty for the battery charge being too far from a set point. Additionally, the total fuel consumed and total emissions generated while driving are measured and penalized. Ideally, the optimization problem posed for designing a controller should minimize some combination of the total fuel consumption, the total emissions generated and the deviation of the SOC from a set point when the vehicle is turned off. Towards this goal, this paper investigates the use of Shortest Path Stochastic Dynamic Programming (SP-SDP) to design an optimal control law.

SP-SDP is a variation of infinite horizon Stochastic Dynamic Programming. In SP-SDP, the state space is constructed so there is a set of states that represent a terminal condition. In this terminal condition, no costs are incurred and there is zero probability of transitioning out of this set. The terminal set is absorbing. Every element of the state space outside of the terminal set is constructed so that, under control, there is a positive probability of transitioning into the terminal set within a finite time. The existence of this terminal set and the ability to control the system, so that the state transitions into the terminal set, guarantees that the expected costs in the system are finite, even in the undiscounted case.

2) Models

2.1) An HEV Dynamic Model.

This study assumes the HEV model of [15] and [20]. A detailed description of the model is available at [21]. The vehicle is a parallel hybrid electric truck, with an automatic transmission and diesel engine. The electric machine is coupled to the output shaft of the transmission. Figure 1 provides a schematic of the HEV powertrain and controls. For control design, the HEV powertrain is simplified. The vehicle is reduced to a quasi-steady, discrete time, state model with two state variables and two inputs. There is a state associated with the charge in the battery (q) and a state associated with the vehicle speed (v). The charge is normalized to a value between 0 and 1. The vehicle speed is limited to a range of 0 to 100 miles per hour. The state variables are assumed to be perfectly observed. The battery is modeled using ohmic resistance and charge efficiency with an open-circuit voltage curve that is a function of the state of charge. The transmission gear

control is simplified to a function of engine speed, engine torque and vehicle speed. The transmission gear is selected apriori to minimize the instantaneous linear combination of fuel consumption and emissions generation that matches the weighting in the control objectives. Gear shifts are assumed to occur in negligible time and their control is provided by an autonomous low-level controller. The vehicle command from the driver, V_{cmd} , is either a wheel power demand, a command to turn the vehicle off or an indication that the vehicle is off where

$$V_{cmd} \in \left\{ \left[P_{\min}, P_{\max} \right], \text{'turn off '}, \text{'off '} \right\}. \tag{1}$$

Power demands are limited by $P_{\min}$ and $P_{\max}$ to a range of -100kW to 150kW as measured at the wheels. When the vehicle command is 'turn off', at the next time step, the vehicle command transitions to 'off'. Once the vehicle command transitions to 'off', the vehicle command stays at 'off'. The deterministic dynamics associated with the HEV are time invariant and summarized as

$$\begin{bmatrix} q & v \end{bmatrix}_{k+1}^{T} = f_{HEV}\left(\begin{bmatrix} q & v \end{bmatrix}_{k}^{T}, \begin{bmatrix} PSR & V_{cmd} \end{bmatrix}_{k}^{T} \right). \tag{2}$$

The deterministic HEV model also generates outputs associated with instantaneous fuel consumption and emissions. When the vehicle command, V_{cmd} , is 'turn off' or 'off', these outputs are 0. These relationships are time invariant and summarized as

$$\begin{bmatrix} \dot{m}_f & \dot{m}_{PM} & \dot{m}_{NOX} \end{bmatrix}_{k}^{T} = h_{HEV}\left(\begin{bmatrix} q & v \end{bmatrix}_{k}^{T}, \begin{bmatrix} PSR & T_{dem} \end{bmatrix}_{k}^{T} \right), \tag{3}$$

where $\dot{m}_f$ is engine fueling rate, and $\dot{m}_{NOx}$ and $\dot{m}_{PM}$ are respectively NOx and particulate matter emission from the engine. The inputs to the HEV are the vehicle command from the driver, V_{cmd} and the power split ratio from the supervisory controller, PSR , which is defined as

$$PSR = \frac{\text{Engine Power Measured at the Wheels}}{\text{Power Measured at the Wheels}}. \tag{4}$$

There exists a low-level controller that interprets a Power Split Ratio (PSR) command and where feasible, delivers engine power, measured at the wheels, that is equal to the product of PSR and the power demand in the vehicle command. Inertial effects are ignored in calculating this ratio. Gearing and losses are considered

in this equation. Any power excess or deficit is balanced by the electric machine. Figure 2 illustrates the signal flow for this system.

2.2) Forming a Controlled Markov Process from a Conditional Probability Drive Cycle Model.

One way to model a vehicle drive cycle is as a white noise process that is transformed into power demands by the environment and driver. This transformation can be visualized as shown in Figure 2. A white noise describes the unique changes in the environment as the vehicle moves through it. The environment correlates these changes and in turn, provides cues to the driver to adjust his internal state. Based on the immediate environmental cues, feedback on vehicle position and velocity and the driver's internal state, the desired wheel torque is determined. This is a complex process and multiple approaches have been used in modeling.

A stationary Markov model based on 'microtrips' is proposed in [19]. The 'micro trips' describe maneuvers over a few seconds: accelerations, quasi-cruising, cruising and decelerations. A collection of 'micro trips' is created using statistical analysis of representative driving data. A probability transition matrix describes the conditional probability of transitioning from one 'micro trip' to another.

Another way to characterize a drive cycle is using conditional probabilities coupled with deterministic models, as described in [20]. This results in a compact description of a stationary controlled Markov process **[22]**. Because the concept of 'turning the vehicle off' is included in this model, the vehicle command can either be a power level, a command to turn the vehicle off, or a command to keep the vehicle off. A conditional probability model is created that defines the probability of the next vehicle command given the current vehicle speed and the current vehicle command. The values of the transition probabilities are calculated from representative data that includes speed and power demands. This conditional probability, for the drive cycle, combined with the HEV dynamic model, forms a controlled Markov process. The state space is the 3-tuple (vehicle command, velocity, battery charge). The controllable input is the power split ratio. The resulting conditional probability of the next vehicle command is

$$V_{cmd,next} \sim \Pr\left(V_{cmd,next} = W \left| V_{cmd}, v \right.\right). \tag{5}$$

Once the current vehicle command is 'off', all subsequent vehicle commands are 'off'.

Wheel power is used to represent the driver's demand because the range of driver requests is primarily limited by engine and machine power. When using a rectangular grid of fixed resolution for approximating (5), the use of power rather then torque results in more points with non-zero probabilities. This is advantageous because it increases the amount of information available versus torque when using simple sampling schemes. One issue with using power rather than torque is how to handle a velocity of zero. Considering only positive speeds, at approximately zero velocity, the power is found by assuming that vehicle speed is a small positive number. To implement this, the driver's request in terms of torque is converted to power using

$$P_{whl} = \begin{cases} T_{whl} \cdot v & , |v| > \varepsilon \\ T_{whl} \cdot \varepsilon & , \text{otherwise} \end{cases}, \tag{6}$$

where ε is 1 mile per hour.

3) Control Law Formulations

3.1) Optimal Cycle-Specific Formulation Using Deterministic Dynamic Programming.

For reference, the optimal cycle-specific controller is illustrated. This is introduced to highlight the differences between the deterministic programming approach and the stochastic dynamic programming approach. The goal of the deterministic controller is to minimize weighted fuel and emissions over a finite horizon for a specific cycle, while returning the battery to a know state of charge when the vehicle is turned off. The optimization problem is formulated as

$$J = \min_{PSR_k \in \mathcal{U}_k(q_k)} \left\{ \sum_{k=1}^{K-1} c_{inst}^T \cdot h_{HEV,k}\left(q_k, PSR_k\right) \right\}, \tag{7}$$

subject to the system dynamics,

$$q_{k+1} = f_{HEV,k}\left(q_k, PSR_k\right), \tag{8}$$

and a requirement to balance the charge in the battery at the end of the cycle,

$$q_K = q_1 . \tag{9}$$

In this problem, $f_{HEV,k}$, is the time-varying dynamics formed by a single trajectory or sample path of $V_{cmd,k}$ and eliminating that term from equation (2). The time-varying fuel and emissions of the system,

$h_{HEV,k}$, are formed from (3) in a similar manner. The product of c_{inst}^T and $h_{HEV,k}$ determines the instantaneous cost of operation of the HEV. This weighting allows the fuel economy to be traded off relative to the emissions. The set $\mathcal{U}_k$ contains all of the feasible choices for the power split ratio at time step k.

Under the assumption of perfect knowledge of the trajectory of the system, an optimal open-loop control law can be found. However, this formulation of the optimal control problem is problematic because there are states at specific time steps that cannot satisfy (9). For example, if the battery is fully charged at K-1 and q_1 is 50% charge, there may be no action that can satisfy (9). Because of these infeasible states, either the constraint formulation must be changed or auxiliary control objectives must be defined to handle infeasible regions. By revising the objective in (7) to include a quadratic cost when turning the vehicle off, the equality constraint in (9) can be approximated. This revised objective is

$$J = \min_{PSR_k \in \mathcal{U}_k(q_k)} \left\{ \left(\sum_{k=1}^{K-1} c_{inst}^T \cdot h_{HEV,k}\left(q_k, PSR_k\right) \right) + \cdot c_{term} \cdot \left(q_K - q_{\text{target}}\right)^2 \right\}. \quad (10)$$

In this form, a time-varying feedback control law can be constructed that is valid for any initial condition. In (10), the constant c_{term} is used to scale a quadratic penalty for deviation from a battery state of charge setpoint at termination of the drive cycle.

A time varying feedback controller for (10) can be found by forming a DP and solving for the value function [23], [24], [25]

$$V_k^*(q) = \begin{cases} \min_{PSR \in \mathcal{U}_k(q)} \left(c_{inst}^T \cdot h_{HEV,k}\left(q, PSR\right) + V_{k+1}^*\left(f_{HEV.k}\left(q, PSR\right)\right)\right) & \text{, if } k \in \{1,...,K-1\} \\ \min_{PSR \in \mathcal{U}_k(q)} \left(c_{term} \cdot \left(q_K - q_{\text{target}}\right)^2 \right) & \text{, if } k = K \end{cases} . \quad (11)$$

Once the value functions has been computed, the optional control law for the power split ratio is determined by solving

$$PSR_k^*(q) \in \arg\min_{PSR} \left(c_{inst}^T \cdot h_{HEV,k}\left(q, PSR\right) + V_{k+1}^*\left(f_{HEV.k}\left(q, PSR\right)\right)\right), \forall k \in \{1,...,K-1\} . \quad (12)$$

Similar formulations have been presented in [13], [14] and [15]. The use finite horizon dynamic programming to design a time varying feedback control law has three distinct drawbacks. First, the optimality of the control does not apply if the drive cycle is different than the one used in (11). Second, the optimal action is not defined after time step K. Third, for implementation in a vehicle, this control law can be considered a 'cycle beater': a strategy that is specifically tuned to perform well on a testing regime while behaving significantly differently in other circumstances. 'Cycle beater' strategies are not permitted by the regulatory agencies that certify vehicles. For this reason, among others, the work in [15] and [20] focused on stationary (i.e. time-invariant) control laws and assumed that perfect state information was available. The restriction to stationary control laws is a tradeoff. It allows the control law to avoid the 'cycle beating' designation. However, in general, it reduces the optimality of the resulting control laws.

3.2) The Ideal SP-SDP Problem Formulation for a Stationary Markov process.

The focus of this paper is on the formulation of the HEV control problem as a Shortest Path Stochastic Dynamic Program (SP-SDP). This formulation offers two advantages over existing formulations. It allows undiscounted future costs and assignment of a cost that is incurred only when the system enters a 'terminal' state. To guarantee that a solution to the SP-SDP exists, the 'terminal' state must be absorbing and all other states transient. Absorbing mean that, once the system enters the terminal state, it stays there, and transient means that every state has a non-zero probability of entering and the absorbing state in finite time. While in the terminal state, no costs are incurred. The SP-SDP formulation is advantageous because it does not require a discount factor and it allows costs to be described based on the fuel, emissions and SOC objectives specific to vehicle certification. Additionally, this formulation reduces the number of tuning parameters used in previous work. The SP-SDP formulation works well with vehicle control problems since there is a guaranteed termination of the drive cycle when the key is turned off. This controller minimizes

$$J = \min_{PSR_k \in \mathcal{U}(x_k)} \left\{ E^W \left(\sum_{k=1}^{\infty} \gamma^k \cdot \begin{cases} c_{inst}^T \cdot h_{hev}\left(q, v, V_{cmd}, PSR\right) & , V_{cmd} \notin \{\text{'off','turning off'}\} \\ c_{term} \cdot \left(q_K - q_{\text{target}}\right)^2 & , V_{cmd} = \{\text{'turning off'}\} \\ 0 & , V_{cmd} = \{\text{'off'}\} \end{cases} \right) \right\} \quad (13)$$

subject to

$$q_{k+1} = f_{HEV}\left(q_k, PSR_k\right), \tag{14}$$

and (5). The constant γ is introduced as a discount factor to keep the form of the equations similar to the infinite horizon controller. In the design of the SP-SDP controller, the discount factor is 1.

By manipulation of this formulation, a SP-SDP is created. The stochastic model in (5) is augmented with the dynamics in (14) to form a controlled Markov decision process. The stochastic disturbance to the system is

$$w = \left[V_{cmd,next}\right]. \tag{15}$$

The subscript 'next' is used to indicate that this is the next vehicle command to be selected by the driver. The controllable input is the power split ratio

$$u = \left[PSR\right]. \tag{16}$$

The power split ratio is restricted to the set, $\mathcal{U}\left(x\right)$, of power splits that are capable of meeting the driver's vehicle command:

$$\mathcal{U}\left(x\right) = \left\{\text{set of all power split ratios that can meet } V_{cmd} \text{ for state } x\right\}. \tag{17}$$

Because of the sizing of the engine and motor, for this class of problems, this set is always non-empty. The system's state is augmented to include the current vehicle command and defined as

$$x = \left[q \quad v \quad V_{cmd}\right]^T. \tag{18}$$

The state propagation equation from (2) is augmented to include memory of the next vehicle command resulting in a new state propagation equation

$$x_{k+1} = f_{HEV,aug}\left(x_k, u_k, w_k\right) = \begin{bmatrix} f_{HEV}\left(\left[q \quad v\right]_k^T, \left[PSR \quad V_{cmd}\right]_k\right) \\ V_{cmd,next,k} \end{bmatrix}. \tag{19}$$

As in (10), the instantaneous cost of operation is a weighted sum of fuel consumption and emissions. If the vehicle is 'off', then no additional costs are accrued. For best fuel economy, c_{inst} is equal to $\left[1 \quad 0 \quad 0\right]^T$. Following the convention in [15], to balance fuel economy and emissions optimization, c_{inst} is equal to

$\begin{bmatrix} 1 & \mu & \upsilon \end{bmatrix}^T$, where μ is the weight associated with NOx emissions and υ is the weight associated with particulate matter emissions. The constant c_{term} is nominally large and its value is selected by trial and error to obtain a battery state of charge at the end of the cycle that is acceptable. In this paper, c_{term} = 1,000,000 is used. The resulting cost function is

$$c(x,u,w)=\begin{cases} c_{inst}^T \cdot h_{HEV}\left(\begin{bmatrix} q & v \end{bmatrix}^T, \begin{bmatrix} PSR & V_{cmd} \end{bmatrix}\right) & \text{,if } V_{cmd} \notin \{\text{'off','turning off'}\} \\ c_{term} \cdot \left(q - q_{\text{target}}\right)^2 & \text{,if } V_{cmd} = \{\text{'turning off'}\} \\ 0 & \text{,if } V_{cmd} = \{\text{'off'}\} \end{cases}. \quad (20)$$

The system model and cost functions are combined to solve for the value function

$$V^*(x)=\min_{u\in\mathcal{U}(x)}\left(E_W\left(c(x,u,w)+\gamma\cdot V^*\left(f_{HEV,aug}(x,u,w)\right)\right)\right). \quad (21)$$

In this equation, the discount factor γ is introduced for eas of comparison with discounted-cost control laws, and is set to one for the SP-SDP formulation. This value function is used to find the set of optimal torque split values via

$$PSR^*(x)\in\arg\min_{u\in\mathcal{U}(x)}\left(E_W\left(c(x,u,w)+\gamma\cdot V^*\left(f_{HEV,aug}(x,u,w)\right)\right)\right). \quad (22)$$

Since there is no requirement for continuity in sequences of power splits, the choice from this set is arbitrary. In the case of multiple elements, the first element is selected. Note, when realized for vehicle control, the vehicle command in the state variable, x, is based on the driver's vehicle command at the instant of evaluation.

It would be desirable to solve this problem for the continuous state and action spaces in which it was posed. However, computing a solution is difficult since it involves finding elements in infinite dimensional spaces. These are not easily found. The generation of a continuous state Markov process is also difficult. Because of this, the problem is solved approximately. The SP-SDP is formulated over a discrete state, action and noise space. The approximate value function, for this discrete problem, is found numerically. The approximate value function is then used to determine the control law for a continuous range of states by interpolation.

3.3) The Approximate SP-SDP Problem Formulation.

To approximate the ideal SP-SDP in a form that is numerically tractable, the state space, action space and noise space are quantized as finite sets. The conditional probability model in Section 2.2 is combined with the deterministic dynamics in section 2.1 to form a controlled Markov process. The cost functions for the SP-SDP are defined as in (20).

3.3.1) Finding the Conditional Probability Model of Driving

To generate a conditional probability model of the vehicle commands, (5), the wheel power and vehicle speed are quantitized over a family of drive cycles. To distinguish quantitized forms of otherwise continuous variables, the crescent $\breve{(\bullet)}$ will be placed over those variables. The power demand is quantitized into N_P discrete levels. The discrete power demand, $\breve{P}_{dem}$ takes values in the set of power demand $\breve{\mathcal{P}}$, where

$$\breve{P}_{dem} \in \breve{\mathcal{P}} = \left\{ P_{dem}^1, P_{dem}^2, \ldots, P_{dem}^{N_P} \right\}. \tag{23}$$

The vehicle command, $\breve{V}_{cmd}$, is selected from the set of vehicle commands, $\mathcal{C}$, which is the union of the power demands with the 'turn off' and 'off' states,

$$\breve{V}_{cmd} \in \breve{\mathcal{C}} = \left\{ \text{'off '}, \text{'turn off '} \right\} \bigcup \breve{\mathcal{P}}. \tag{24}$$

The vehicle velocity is quantized into N_v discrete levels. The velocity of the vehicle is assigned to an element in the set of possible velocity values $\breve{\mathcal{V}}$,

$$\breve{v} \in \breve{\mathcal{V}} = \left\{ v_{wh}^1, v_{wh}^2, \ldots, v_{wh}^{N_v} \right\}. \tag{25}$$

To determine the values of $\Pr\left\{ \breve{V}_{dem,k+1} \middle| \breve{V}_{dem,k}, \breve{v}_k \right\}$, a maximum likelihood method similar to [15] is used. The conditional probability is calculated by first selecting a set of drive cycles and using the vehicle model to obtain wheel power requests needed to follow the drive cycle. The vehicle speed versus time is part of the drive cycle specification. Next, this data is sampled at a 0.05 second interval, and formed into a list of 3-tuples of data: $\left(P_{dem}\left(t_k + 1\right), P_{dem}\left(t_k\right), V\left(t_k\right) \right)$ where $t_k \in \left\{ 0, 0.05, 0.10, \ldots \right\}$. These 3-tuples of data

are quantized using a nearest neighbor method and assigned to elements in (23) and (25). The resulting conditional probability model is formed as

$$\Pr\left\{\breve{V}_{cmd,next}\left|\breve{V}_{cmd},\breve{v}\right.\right\}\triangleq\frac{\left\{\begin{array}{c}\text{Total number of samples matching } \breve{V}_{cmd,next} \text{ and } \breve{V}_{cmd} \text{ and } \breve{v}\\ \text{and vehicle is not 'off' in next sample}\end{array}\right\}}{\text{Total number of samples matching } \breve{V}_{cmd} \text{ and } \breve{v}}. \quad (26)$$

$$,\forall\breve{V}_{cmd}\in\breve{\rho},\breve{V}_{cmd,next}\in\breve{\rho} \quad (27)$$

This conditional probability model has the property that for some speeds and power demands, the sum of probabilities for the next power demand is less than 1. This difference is the probability that at the next time step, the vehicle command will be 'turn off' where

$$\Pr\left\{\text{'turn off '}\left|\breve{V}_{cmd},\breve{v}\right.\right\}\triangleq 1-\sum_{\breve{P}_{dem,next}\in\breve{\rho}}\Pr\left\{\breve{P}_{dem,next}\left|\breve{V}_{cmd},\breve{v}_k\right.\right\},\forall\breve{V}_{cmd}\in\breve{\rho}\,. \quad (28)$$

Further, the probability of the vehicle being 'off' is given by

$$\Pr\left\{\text{'off '}\left|\breve{V}_{cmd},\breve{v}\right.\right\}\triangleq\begin{cases}1 & ,\text{if } \left(\breve{V}_{cmd}=\text{'turn off'}\right) \text{ or } \left(\breve{V}_{cmd}=\text{'off'}\right)\\ 0 & ,\text{otherwise}\end{cases}. \quad (29)$$

Depending on the data set used to calculate the conditional probability, there may not be sufficient data available to guarantee that the denominator in (26) is nonzero. The following rules were used to guarantee that the probability distribution is defined for all points. For each triple of next power demand, current power demand and velocity tuple where the denominator in (26) is zero, if the current power demand is greater than 0, the next power demand is set to the next lowest power demand in the set $\breve{\rho}$. If the current power demand is less than 0 and the velocity is greater than 0, the next power demand is set to the minimum of the next greater power demand and the greatest negative element in $\breve{\rho}$. If the current power demand is less than 0 and the vehicle speed is 0, the next power demand is set to 0. If the current power demand is 0 and the velocity is greater than 0, the next power demand will be the negative element in $\breve{\rho}$ closest to 0. If the current power demand is 0 and the velocity is 0, the next power demand is set to 0.

3.3.2) Interpolation for Solution of the SP-SDP

In the continuous formulation of the DP problem, the state propagates from one distinct state to another. When the state is quantized, a different approach is required. The propagation is modeled via partial membership in multiple discrete states. This is equivalent to modeling the propagation stochastically, where the partial membership in the multiple states is equal to the probability of transition into any single discrete state.

For this quantization, a membership function is used that maps every value in the continuous state space to membership in discrete state space per

$$M_{\mathcal{X},\breve{\mathcal{X}}} : \mathcal{X} \to [0,1]^{|\breve{\mathcal{X}}|}. \tag{30}$$

This function maps a continuous space $\mathcal{X}$ to a vector of values between 0 and 1 with one element for each element in the discrete set $\breve{\mathcal{X}}$. In performing this quantization, there are several choices in how the quantization is performed and how the membership function is implemented. The quantization can be performed on a regular or irregular grid. For this work, an irregularly spaced rectangular grid was used. The set of points for this grid are denoted as $\breve{\mathcal{X}}$. The membership function can be implemented using several different types of interpolation schemes. Different schemes have different tradeoffs to consider in terms of accuracy and complexity [26],[27],[28],[29],[30],[31]. Two methods were investigated: bilinear [26] and barycentric interpolation [30]. Barycentric interpolation was found to provide a better trade off between accuracy and complexity than bilinear interpolation.

3.3.3) The Discrete SP-SDP Equations.

Using the definitions and functions developed in Sections 3.3.1 and 3.3.2, the SP-SDP is formulated. The state space is quantized into the discrete set of points in the set $\breve{\mathcal{X}}$. The power split ratios are quantized into a discrete set $\breve{\mathcal{U}}$. This discrete set is restricted to ensure that the choices can best meet the vehicle command at the current state. This restricted set is represented as $\mathcal{U}(x)$. The next vehicle commands are quantized as in (23). The continuous state SP-SDP equations (21) are adapted to obtain the discrete state SP-SDP equations

$$V^*\left(\breve{x}\right)=\min_{\breve{u}\in\breve{U}(x)}\left(\sum_{\breve{w}\in\breve{\mathcal{C}}}\Pr\left(\breve{w}|\breve{x}\right)\left(c\left(\breve{x},\breve{u},\breve{w}\right)+\gamma\cdot\left(V^*\right)^T\cdot M_{\breve{\chi},\breve{\chi}}\left(f_{HEV,aug}\left(\breve{x},\breve{u},\breve{w}\right)\right)\right)\right),\forall\breve{x}\in\breve{\chi}\ .(31)$$

Since the state is quantized, the value function is a vector, with a single value that corresponds to each discrete state. The optimal control law for the power split ratio is then found using the value function from (31) in

$$PSR^*\left(\breve{x}\right)\in\arg\min_{u\in\breve{\mathcal{U}}(\breve{x})}\left(\sum_{\breve{w}\in\breve{\mathcal{C}}}\Pr\left(\breve{w}|\breve{x}\right)\left(c\left(\breve{x},\breve{u},\breve{w}\right)+\gamma\cdot\left(V^*\right)^T\cdot M_{\breve{\chi},\breve{\chi}}\left(f_{HEV,aug}\left(\breve{x},\breve{u},\breve{w}\right)\right)\right)\right),\forall\breve{x}\in\breve{\chi}\ .(32)$$

In the case where more than one power split satisfies (32), the lowest value element is chosen.

3.3.4) Interpretation of the Value Function

The value function, $V^*\left(x\right)$, that results from solving the SP-SDP is the expected cost of operation of the HEV from the state x. Since the HEV will always start a drive cycle from the 0-speed, 0-torque request point, the value function at this point provides a measure of the expected costs that will be incurred over the family of drive cycles described by the conditional probability model in (26) through (29).

3.4) Numerical Solution.

To solve for the value function, a Linear Program (LP) was formulated as discussed in [23]. This was done, rather than using Value Iteration (VI) or Policy Iteration (PI), for two reasons. First, by creating an LP, there was no need to generate codes for VI or PI. This was a significant savings in engineering effort. Secondly, LPs can be solved very efficiently due to their convexity and the existence of efficient commercial codes.

The LP can be formed because of the discrete state and action spaces used in the approximate problem formulation. The general idea in forming a linear program is that a single scalar value represents the value function for each discrete state. This single value must be equal to the minimum obtained when all actions are considered. By generating an inequality for each discrete action at any discrete state, the constraints on the value function can be formed. Consider the discrete dynamic programming equations where the action u is restricted to a discrete action space, the state space is discrete and the evolution of the system is restricted to those discrete states. The minimization statement,

$$V(x) = \min_{u \in \mathcal{U}} \left(c(x,u) + \gamma \cdot V\left(f(x,u) \right) \right), \tag{33}$$

implies that, when $\mathcal{U}$ is a finite and countable set, $V(x)$ satisfies a finite system of equations

$$\left\{ \begin{array}{l} V(x) \le \left(c(x,u_1) + \gamma \cdot V\left(f(x,u_1) \right) \right) \\ V(x) \le \left(c(x,u_2) + \gamma \cdot V\left(f(x,u_2) \right) \right) \\ \vdots \\ V(x) \le \left(c(x,u_N) + \gamma \cdot V\left(f(x,u_N) \right) \right) \end{array} \right\}, \tag{34}$$

and will satisfy at least one of those equations as an equality.

Furthermore, the set of inequalities in (33) is extended to apply to each discrete state in the DP equations. For each of these sets of inequalities, there is a single scalar value that represents the value function at that discrete state. By maximizing the sum of the value function at each discrete state, an optimization problem is formed. Since the cost at each combination of state and actions is known, these become constants in the problem statement. Moreover, the value function is represented as a real vector. If the value function on the right hand side of (33) is moved to the left hand side, the inequality can be restated as a linear combination of value functions at specific discrete states being less than a constant as

$$\left\{ \begin{array}{l} V(x) - \gamma \cdot V\left(f(x,u_1) \right) \le c(x,u_1) \\ V(x) - \gamma \cdot V\left(f(x,u_2) \right) \le c(x,u_2) \\ \vdots \\ V(x) - \gamma \cdot V\left(f(x,u_N) \right) \le c(x,u_N) \end{array} \right\}, \forall x \in \breve{\mathcal{X}} \Rightarrow \left\{ \begin{array}{l} A \cdot V \le b \\ V \in \mathbb{R}^{|\breve{\mathcal{X}}|} \\ A \in \mathbb{R}^{|\breve{\mathcal{X}}| \cdot |\breve{\mathcal{U}}| \times |\breve{\mathcal{X}}|} \\ b \in \mathbb{R}^{|\breve{\mathcal{X}}| \cdot |\breve{\mathcal{U}}|} \end{array} \right\}. \tag{35}$$

This set of inequalities forms linear constraints on the value function. By constructing an optimization problem to maximize the sum of the elements of V subject to the inequalities in (35), a linear program is formed that solves the dynamic programming equations.

A similar procedure was used to convert the SP-SDP equations in (31) into a matrix and vector defining the constraints in a linear program.

4) Control Design Results.

The controller obtained using stochastic dynamic programming is a function that maps vehicle speed, battery SOC and the vehicle command to a power split ratio. For evaluation, the SP-SDP controller and an infinite-horizon, discounted-cost controller designed via SDP are compared. The infinite-horizon, discounted-cost controller is referred to as the SDP controller; it is based on the controllers presented in **[20**] and **[32**]. The SDP controller minimizes

$$J = \min_{PSR_k \in \mathcal{U}(x_k)} \left\{ E^W \left(\sum_{k=1}^{\infty} \gamma^k \cdot \left(c_{inst}^T \cdot h_{hev}\left(q, v, V_{cmd}, PSR\right) + K_q \cdot \left(q_K - q_{\text{target}}\right)^2 \right) \right) \right\}, \qquad (36)$$

subject to (14) and (5). A discount factor of 0.95 is used to match the work in **[32**]. The probability model, (5) is modified so that the probability of 'turning off' is 0. For the SDP controller K_q is set to 60,000.

Both the SDP and the SP-SDP controllers use the same value for c_{inst}. The weighting is one times the fuel rate in grams per second plus forty times the NOx emissions rate in grams per second plus eight hundred time the particulate matter emission rate in grams per second. In both cases, q_{target} is 0.5.

For both the SDP and SP-SDP cases, four different controllers were designed. These controllers were designed using the UDDS Heavy Duty Cycle, the FET Highway Cycle, the WVU Suburban Cycle and the WVU City Cycle as the basis for determining the conditional probability model of the driving pattern. Once the value function was found using (31) and linear programming, a higher resolution grid was applied to the problem and the power split ratio was found using (32). Both controllers were designed using identical quantization and interpolation. The codes to design the controllers were identical except for the equations used to calculate the instantaneous costs, the discount factor and the conditional probability models used to describe the driving pattern.

The resulting control laws were then simulated over the sample path corresponding to the deterministic drive cycles listed above. To best visualize the behavior of the control laws, the engine power as a function of vehicle command and vehicle speed is plotted for the SP-SDP controller in Figure 3 and the SDP controller in Figure 4. These maps are from the controller designed for operation with driving patterns similar to the FET Highway cycle. There are two notable differences to observe in these maps. First the SDP

controller is more aggressive in disabling the engine to improve performance than the SP-SDP controller. Consider the behavior at 10 mile per hour. At this point, the SDP controller disables the engine for battery SOC's greater than about 55% and wheel powers below about 25 kW. Conversely, the SP-SDP controller only disables the engine above 65% battery SOC and below about 25 kW wheel power. Additionally, the SDP controller demonstrates significantly more use of engine power to control battery SOC at all speeds than the SP-SDP controller does. Consider the behavior at 10 miles per hour. At this speed, the SDP controller has the engine running to the charging limit of the battery for SOC values below about 40%. Conversely, for the same operating point, the SP-SDP controller uses less engine power for the same operating points.

5) Comparison and Simulation Results.

Comparison to previously published results is difficult since the SOC values for equilibrium operation are not available. However, this work showed the improvement resulting from SDP versus Heuristic methods for improvement in fuel economy and emissions. It follows that the relative performance obtained there holds for these results.

All of these controllers were evaluated by simulation. The only difference in the simulations is the selection of the drive cycle and the selection of the control law, based on power split ratios, used. The performance of each controller is found by executing the model such that the starting battery SOC is equal to the final battery SOC within a tolerance of 0.0001. Under these conditions, the performance is measured as the sum of the total fuel consumed plus forty times the total Nitrous Oxides emitted plus eight hundred times the total particulate matter emitted divided by the total distance in miles. For the performance measure, smaller numbers are better. An example of the SOC under the SP-SDP controller for the HWFET drive cycle is shown in Figure 5. It should be noted that the changes in SOC only require a small fraction of the energy capacity of the battery. For this vehicle, the battery is sized for durability which results in a battery with significantly more energy capacity than that needed for energy swings alone. The performance results of the SP-SDP controller are summarized in Table 1. This table shows the performance of the SP-SDP controller trained against a single cycle controlling the vehicle over a suite of drive cycles.

Table 1 - SP-SDP Controller Performance

	Test Cycle								
Training Cycle	UDDS HDV	WVU Interstate	WVU Suburban	WVU City	Highway FET	SC03	NYC Composite	NYC Truck	Manhattan
UDDS HDV	833.58	1031.80	756.99	576.91	914.22	1265.90	400.81	708.73	1170.70
WVU Suburban	849.37	908.07	627.05	474.20	985.62	1222.30	366.65	678.09	1298.46
WVU City	889.35	939.72	731.71	509.46	898.44	1163.20	408.18	852.72	1437.18
FET Highway	938.53	1045.80	561.58	705.48	944.14	1200.50	394.765	821.59	1125.71

For each of the tests, the equilibrium SOC at key off is also summarized in Table 2. The equilibrium SOC is the SOC obtained when the vehicle is driven on a particular cycle repeatedly until the starting and ending SOC are effectively the same value. In this case the cycles are repeated until the difference between the starting and ending SOC is less than 0.0001.

Table 2 - SP-SDP Equilibrium SOC when Vehicle is Turned Off

	Test Cycle								
Training Cycle	UDDS HDV	WVU Interstate	WVU Suburban	WVU City	Highway FET	SC03	NYC Composite	NYC Truck	Manhattan
UDDS HDV	0.4982	0.4626	0.4722	0.5007	0.4774	0.4674	0.5027	0.4995	0.4937
WVU Suburban	0.4634	0.5038	0.5058	0.5044	0.4604	0.4677	0.5024	0.4982	0.4895
WVU City	0.4939	0.4974	0.4759	0.4997	0.4615	0.4843	0.4854	0.4932	0.4895
FET Highway	0.4923	0.4831	0.5308	0.5248	0.5004	0.5038	0.5441	0.5198	0.5065

The performance for the SDP controller is summarized in Table 3.

Table 3 - SDP Controller Performance

	Test Cycle								
Training Cycle	UDDS HDV	WVU Interstate	WVU Suburban	WVU City	Highway FET	SC03	NYC Composite	NYC Truck	Manhattan
UDDS HDV	850.46	1033.50	747.93	617.49	921.96	1298.00	459.28	820.97	1059.10
WVU Suburban	861.74	923.85	654.44	513.82	983.64	1172.70	448.59	814.05	1040.80
WVU City	844.74	943.79	699.00	536.82	983.64	1172.70	448.59	814.05	1040.85
FET Highway	943.03	1091.10	574.41	690.14	972.93	1234.40	406.07	826.66	1013.71

The equilibrium SOC for the SDP control law is summarized in Table 4.

Table 4 - SDP Equilibrium SOC when Vehicle is Turned Off

	Test Cycle								
Training Cycle	UDDS HDV	WVU Interstate	WVU Suburban	WVU City	Highway FET	SC03	NYC Composite	NYC Truck	Manhattan
UDDS HDV	0.5128	0.5027	0.5022	0.5057	0.5107	0.5034	0.5131	0.5136	0.5031
WVU Suburban	0.5027	0.5072	0.5103	0.5074	0.5029	0.5030	0.5100	0.5128	0.5080
WVU City	0.5083	0.5090	0.5065	0.5095	0.5034	0.5131	0.5067	0.5128	0.5075
FET Highway	0.5022	0.5026	0.5090	0.5026	0.5214	0.5025	0.5078	0.5136	0.5081

One interesting aspect of these results is that the best performance for both the SP-SDP and the SDP controller does not necessarily occur when the controllers are trained and evaluated using the same drive cycle. For example, see Table 1 for the WVU city evaluation using a SP-SDP. The best performance occurs when the controller is trained using the WVU suburban drive cycle not the WVU city cycle. A similar result occurs when the SDP controller is evaluated on the Highway FET drive cycle. The best performance is measured when the SDP controller is trained using the UDDS Heavy Duty Cycle. This is due to evaluating the performance of the controller on a single realization of the driving cycle and not on the Markov model of the drive cycle. The controller is designed to minimize the expected cost of operation which implies that on any single realization of the drive cycle, there may exist a better controller.

To compare the performance of the SP-SDP controller and the SDP controller, the performance and equilibrium SOC on training drive cycle is considered. The relative performance improvement is measured as the difference between the two controllers divided by the best performance. The reduction in final SOC error is measured as one minus the minimum SOC error divided by the maximum SOC error. These results are summarized in Table 5. In all cases considered, the SP-SDP controller improved performance and improved equilibrium SOC error.

Table 5 - Comparison of Performance in Control Laws

Cycle	SP-SDP Performance	SP-SDP Final SOC	SDP Performance	SDP Final SOC	SP-SDP Improvement Performance	SP-SDP Improvement Reduction in Final SOC Error
UDDS HDV	833.58	0.4982	850.46	0.5128	2.03%	85.94%
WVU Suburban	627.05	0.5058	654.44	0.5103	4.37%	43.7%
WVU City	509.46	0.4997	536.82	0.5095	5.37%	96.84%
FET Highway	944.14	0.5004	972.93	0.5214	3.05%	98.13%

5) Conclusion

By using Shortest Path Stochastic Dynamic Programming to design HEV controllers, it is possible to realize the advantages of Stochastic Dynamic Programming based controller design with better State of Charge control and fewer parameters to tune. Additionally, the optimization criteria for the controller better resembles the engineering goals by solving the control problem without a discount factor and instantaneous cost associated with State of Charge deviation.

While a specific model has been used in this study, Shortest Path Stochastic Dynamic Programming based control design can be applied to any Hybrid Electric Vehicle where the optimal control objective meets two criteria. First there is a cost associated with the state deviating from a specific value when ‘turned off’. Second, the control objective is to minimize the total cost that is a function of the current state and actions plus a cost associated with the state of the system when turned off. Any number of additional states beyond those described here can be part of the model. Hence, this solution technique can be applied to a broad range of HEV powertrains.

References.

[1] Leone TG, Pozar M. Fuel Economy Benefit of Cylinder Deactivation – Sensitvity to Vehicle Application and Operating Constraints. SAE 2001-01-3591.

[2] Lichti TH. Design of a Continuously Variable Cam Phasing (CVCP) System for Emissions, Fuel Economy and Power Improvement. SAE 982960.

[3] Lancefield T. The Influence of Variable Valve Actuation on the Part Load Fuel Economy of a Modern Light-Duty Diesel Engine. SAE 2003-01-0028.

[4] Singh J, Berger K, Mack P, Piorkowski P, Hogan T, Wong A. General Motors "VTi" Electronic Continuously Variable Transaxle. SAE 2003-01-0594.

[5] Link M, Voss B, Eggert E, Nasdal R. The Automated Shift Transmission (AST) – Possibilities and Limits in Production-Type Vehicles. SAE 2001-01-0881.

[6] Wijetunge RS, Hawley JG, Vaughn ND. Application of Alternative EGR and VGT Stratgies to a Diesel Engine. SAE 2004-01-0899.

[7] Recommended Practice for Measuring the Exhaust Emissions and Fuel Economy of Hybrid Electric Vehicles. SAE J1711, March 1999.

[8] Recommended Practice for Measuring Fuel Economy and Exhaust Emissions of Hybrid Electric and Conventional Heavy Duty Vehicles. SAE J2711, Sept 2002.

[9] California Environmental Protection Agency, California Air Resources Board, California Exhaust Emission Standards and Test Procedures for 2003 and Subsequent Model Zero-Emission Vehicles and, 2001 and Subsequent Model Hybrid Electric Vehicles, in the Passenger Car, Light Duty Truck and Medium Duty Truck Classes, Adopted August 5, 1999.

[10] Zero-Emission Vehicle Standards for 2003 and Subsequent Model Passenger Cars, Light-Duty Trucks, and Medium-Duty Vehicles. California Code of Regulations, Title 13, Section 1962.

[11] Pagnelli G, Ercole G, Brahma A, Guezennec Y, Rizzoni G, General Supervisory Control Policy for the Energy Optimization of Charge Sustaining Hybrid Electric Vehicles. *JSAE Review* 2001; 22:11-518.

[12] Johnson V, Wipke K, Rausen D. HEV Control Strategy for Real Time Optimization of Fuel Economy and Emissions. SAE 2000-01-1543.

[13] Tate E, Boyd S. Towards Finding the Ultimate Limits of Performance for Hybrid Electric Vehicles, SAE 2001-01-3099.

[14] Kirschbaum F, Back M, Hart M. Determination of the Fuel Optimal Trajectory for a Vehicle along a Known Route. 15th Triennial World Congress, Barcelona, Spain, IFAC, 2002.

[15] Lin CC, Peng H, Grizzle JW, Kang JM, Power Management Strategy for a Parallel Hybrid Electric Truck. *IEEE Trans. on Control Systems Technology* Nov 2003; 11 (6): 839 – 849.

[16] C.-C. Lin, H. Peng, J. W. Grizzle, J. Liu, and M. Busdiecker, "Control System Development for an Advanced-Technology Medium-Duty Hybrid Electric Truck," *SAE 2003-01-3369*.

[17] Back M, Simons M, Kirschbaum F, Krebs F. Predictive Control of Drivetrains. 15th Triennial World Congress, Barcelona, Spain, IFAC, 2002.

[18] Kummert M, Andre P, Nicolas J, Optimal Heating in a Passive Solar Commercial Building, Solar Energy Vol 69 (Suppl.), Nos 1-6, pp. 103-116, 2000.

[19] Lin J, Niemeier D. Estimating Regional Air Quality Vehicle Emission Inventories: Constructing Robust Driving Cycles. *Transportation Science* August 2003:37 (3); 330-346.

[20] Lin CC, Peng, H, Grizzle, JW. A Stochastic Control Strategy for Hybrid Electric Vehicles, *American Control Conference* 2004.

[21] Jessy W. Grizzle's Publications, 2006. **http://www.eecs.umich.edu/~grizzle/papers/auto.html**

[22] Grimmett G, Stirzaker D. *Probability and Random Processes, 3rd ed*, Oxford University Press Inc.: New York, 2001.

[23] Bertsekas D. *Dynamic Programming and Optimal Control Vol I & II*. Athena Scientific: Belmont, Massachusetts, 1995.

[24] Bertsekas D, Tsitsiklis J. *Neurodynamic Programming*. Athena Scientific: Belmont, Massachusetts, 1996.

[25] Kumar PR, Variaya P. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice Hall: New Jersey, 1986.

[26] Chow CS, Tsitsiklis JN. An Optimal Multigrid Algorithm for Continuous State Discrete Time Stochastic Control. Proceedings of the 27th Conference on Decision and Control, Austin Texas, December, 1988.

[27] Munos R, Moore A. Variable Resolution Discretization in Optimal Control. Machine Learning, 49, 291-323, 2002.

[28] Bertsekas D. Convergence of Discretization Procedures in Dynamic Programming. IEEE Transactions on Automatic Control, 415-419, June 1975.

[29] Chen VCP, Ruppert D, Shoemaker CA. Applying Experimental Design and Regression Splines to High Dimensional Continuous State Dynamic Programming. Operations Research, Vol 47, No 1, 38-53, Jan-Feb 1999.

[30] Munos R, Moore A. Barycentric Interpolators for Continuous Space and Time Reinforcement Learning. Neural Information Processing Systems, 1998.

[31] Moore A. Variable Resolution Dynamic Programming: Efficiently Learning Action Maps in Multivariate Real-valued State-spaces. Machine Learning: Proceedings of the Eighth International Conference, 1991.

[32] Lin CC. Modeling and Control Strategy Development for Hybrid Vehicles. Ph.D. Thesis. Department of Mechanical Engineering, University of Michigan, Ann Arbor, 2004.

Appendix A – Symbols, Abbreviations and Notation

Notation

$\breve{\bullet}$	The crescent is used to indicate that the variable is a discrete variable (set) formed by quantizing a continuous variable (set) that takes on a finite number of values.
$\mathcal{A}$	Set.
$\breve{\mathcal{A}}$	Countable set.
$\#\left\{\breve{\mathcal{A}}\right\}$	The ordinality of a set. This is the number of elements in a set.
$\# x$	The ordinality of the element x in its respective ordered set. This is a nonempty set that describes the order of an element in an ordered set. For example if $x \in \breve{\mathcal{X}} = \{a,b,c,d,e,d\}$, then if $x = c$, $\# x = 3$. Where is $x = d$, we get $\# x = \{4,6\}$.
$[a,b]$	The closed interval along the real line from a to b.
$\{a_1, a_2, ...\}$	An infinite and countable set.
$\{a_1, a_2, ..., a_n\}$	A finite and countable set.
$\Pr\left\{\mathcal{E}\middle\vert\mathcal{F}\right\}$	Conditional Probability: The probability of event $\mathcal{E}$ given that event $\mathcal{F}$ has occurred.
$\Pr\{\mathcal{E}\}$	Probability: The probability of event $\mathcal{E}$.
~	Used to indicate that the variable on the left hand side is a random variable with a probability distribution described by the right hand side. For example, $X \sim \eta(0,1)$ would be read as "the random variable X has a normal distribution with a mean of 0 and a variance of 1."

List of Abbreviations

A/C	Air Conditioning
DP	Dynamic Program
HEV	Hybrid Electric Vehicle
LP	Linear Program
MPC	Model Predictive Control
PHEV	Parallel Hybrid Electric Vehicle
PSR	Power Split Ratio
SDP	Stochastic Dynamic Programming
SOC	State of Charge
SP- SDP	Shortest Path Stochastic Dynamic Programming
TSR	Torque Split Ratio
VCT	Variable Cam Timing

List of Symbols

$c(\cdot)$	Instantaneous cost. Maps a state and a control action to a real number.
$f(\cdot)$	Generally used to indicate the dynamics of the system.
$g(\cdot)$	Generally used to indicate a feedback control law or policy.
$h(\cdot)$	Generally used to indicate the output of the system.
$M_{\mathcal{X},\breve{\mathcal{X}}}(x)$	Membership function that maps an element in the space $\mathcal{X}$ to its membership in each element in $\breve{\mathcal{X}}$.
K_q	Instantaneous SOC quadratic cost gain
q	State of charge of the battery.
u	Action or input to the system.
$\mathcal{U}$	The action space.
$\breve{\mathcal{U}}$	The discrete action space, a finite, countable set.
v	Vehicle speed.
$V(x)$	Value or 'Cost to go' function. Maps the state space to a real number.
x	The state of the system.
$\mathcal{X}$	The space of possible states of the system.
$\breve{\mathcal{X}}$	The discrete state space, a finite, countable set.
$\gamma \in [0,1]$	Discount factor for future costs

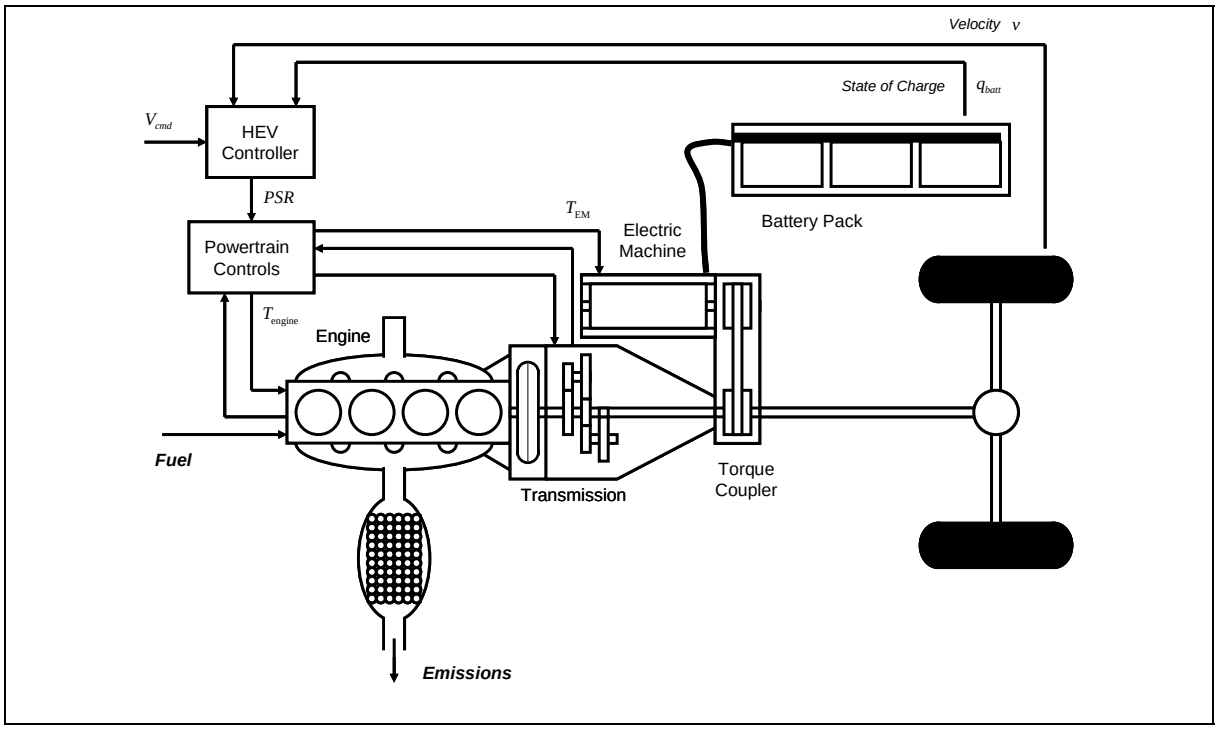

Figure 1 - The Environment, Driver and HEV as a System

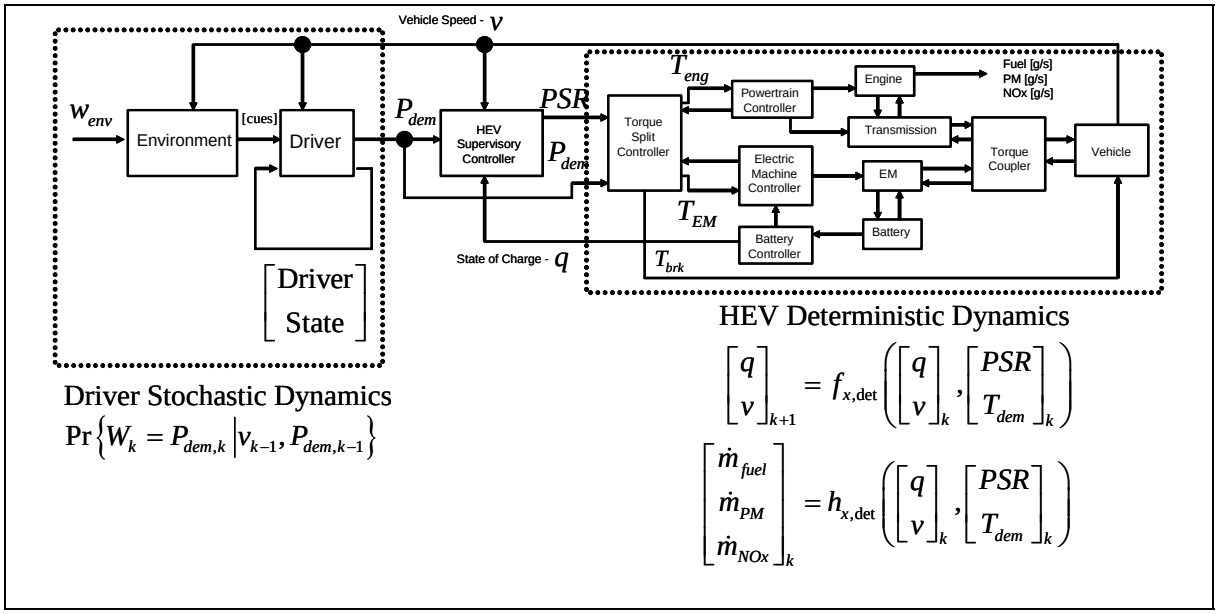

Figure 2 - A Signal Flow Diagram of the Environment, Driver and HEV as a System

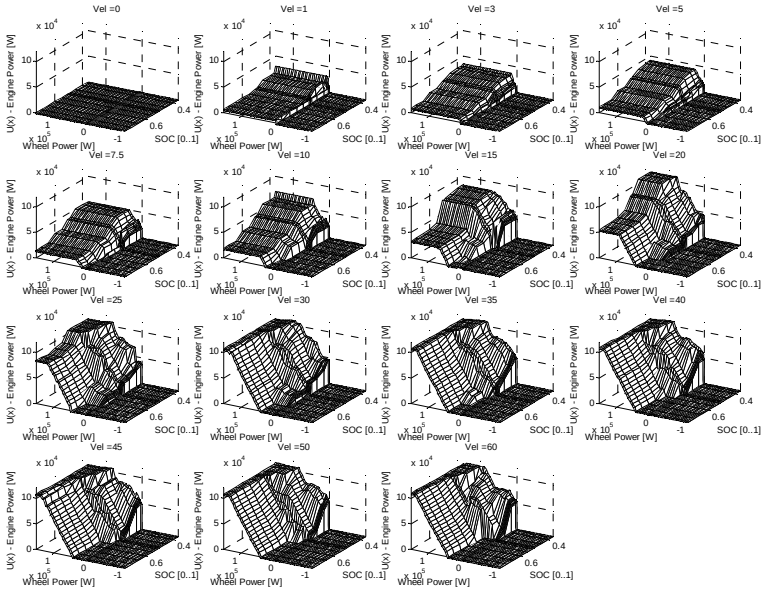

Figure 3 - SP-SDP Controller: Engine Power Map

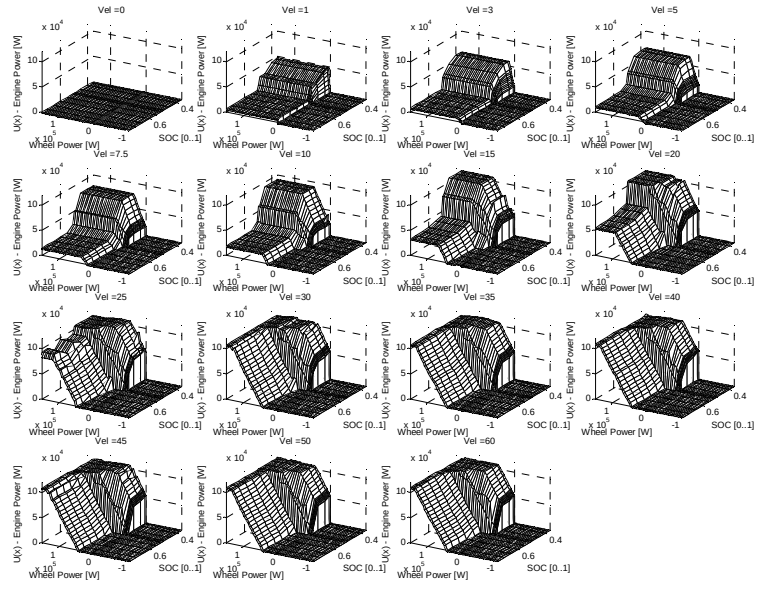

Figure 4 - SDP Controller: Engine Power Map

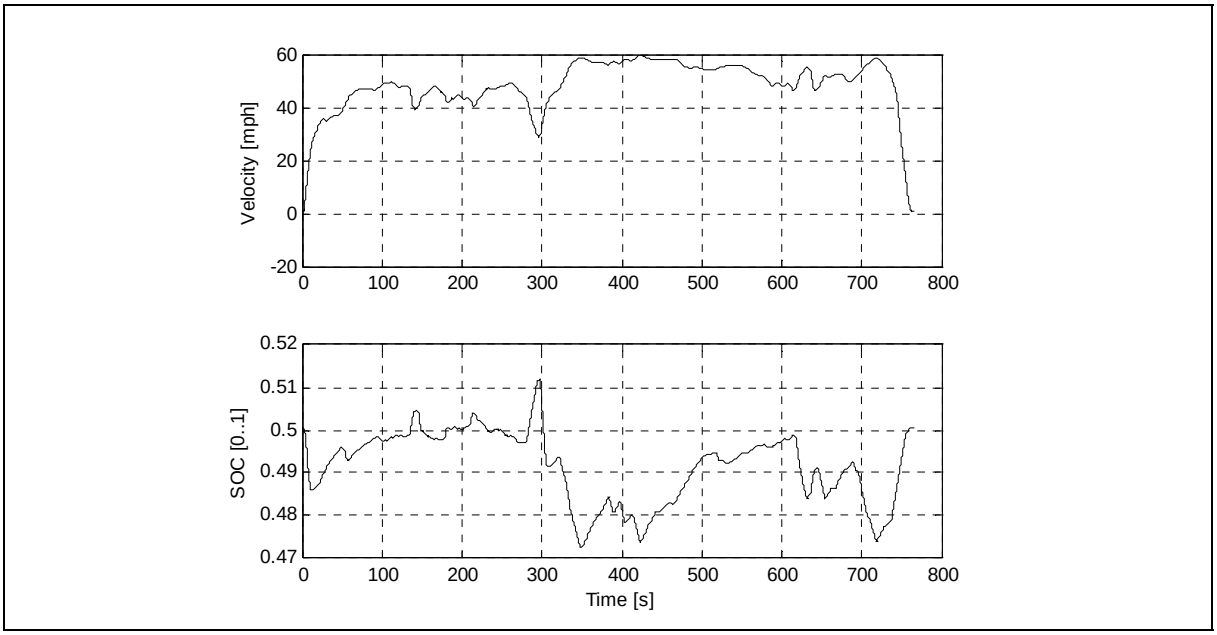

Figure 5 - SOC on the HWFET using the SP-SDP controller