Design and Development of Real-Time Optimal Energy Management System for Hybrid Electric Vehicles

Masood Roohi, Amir Taghavipour

Abstract—This paper describes a strategy to develop an energy management system (EMS) for a charge-sustaining power-split hybrid electric vehicle. This kind of hybrid electric vehicles (HEVs) benefit from the advantages of both parallel and series architecture. However, it gets relatively more complicated to manage power flow between the battery and the engine optimally. The applied strategy in this paper is based on nonlinear model predictive control approach. First of all, an appropriate control-oriented model which was accurate enough and simple was derived. Towards utilization of this controller in real-time, the problem was solved off-line for a vast area of reference signals and initial conditions and stored the computed manipulated variables inside look-up tables. Look-up tables take a little amount of memory. Also, the computational load dramatically decreased, because to find required manipulated variables the controller just needed a simple interpolation between tables.

Keywords—Hybrid electric vehicles, energy management system, nonlinear model predictive control, real-time.

I. INTRODUCTION

FOSSIL fuel powered engines in transportation sector have made some severe consequences such as various lung and heart diseases, greenhouse gas emission, and increasing health expenses for governments. These factors have led the governments to consider strict standards on engine fuel efficiency and emission of the vehicles. One of the short-term approaches of car manufacturers to these new standards and demands of costumers is improving new technologies on sustainable transportation such as (HEVs) hybrid electric vehicles. It is predicted that 90% of consumed energy in transportation will still be provided by fossil fuels in 2030 [1]. Therefore, trying to improve HEVs seems really reasonable.

Generally, HEVs are categorized in three different architectures: series, parallel, and power-split (series-parallel). Despite the complexity of the power-split, it is the most popular architecture among car manufacturers since the power-split architecture can operate in both series and parallel modes. The power-split architecture has schematically shown in Fig. 1.

As it can be seen in Fig. 1, HEVs benefit from two energy converters: internal combustion engine (ICE) and Electrical motor which respectively use fuel and electric as the energy source. Combining these two sources of energy in a vehicle is not a new idea, However, the new generation of HEVs have become more successful than their ancestors because of the advent of new technologies in terms of electronics and control systems [2].

Addition of an energy storage device to the vehicle introduces new flexibility and complexity for the control system. Intelligently utilization and control of this new degree of freedom can lead to fuel economy and decrease emission improvement. The energy management strategy is the high-level control or supervisory control layer uses this flexibility to accomplish the above tasks along with maintaining vehicle drivability [2].

Energy management system controls the power flows from sources to satisfy the control objectives while considering global and local constraints of the power-train. Usually, the primary control objective is the minimization of the vehicle fuel consumption, while minimizing engine emissions and maintaining or enhancing drivability [3]. Towards solving the energy management problem, two general approaches have been introduced. First one is the heuristic approach which has attempted to offer some improvements in the HEV energy efficiency by using expertise. Therefore heuristic approaches may not guarantee either an optimal result in real vehicle operational conditions, or a robust performance if system parameters deviate from their nominal operating points [4]. On the other hand, model-based approaches are inherently more flexible than heuristic approaches and they can fully exploit the potential for energy consumption reduction at the cost of complexity and computational load [5]. Several model-based energy management strategies can be mentioned such as dynamic programming (DP), stochastic dynamic programming (SDP), equivalent fuel consumption minimization strategy (ECMS), and model predictive control (MPC) [6].
MPC can deal with processes with numerous manipulated variables, outputs, and constraints [7]. In addition, MPC is particularly interesting among model-based approaches since it can maintain the robustness of a feedback controller. However, with the limited capabilities of commercial control hardware, the heavy computational cost has become a drawback in practice [8].

The theory of the applied approach is described in Section II. The control-oriented model that is very important in a model-based approach is derived in Section III. Designing the energy management system is explained in Section IV. Simulation results are laid out in Section V followed by the conclusions in Section VI.

II. THEORY

It is possible to define several criteria for a control system such as being fast, suppressing overshoot, considering the limitation on manipulated variables. Any reasonable criterion can be defined to be achieved by the predictive controller. As it was mentioned in the introduction section, there can be several control objectives to be achieved in the energy management problem. In this case, the major control objective is improving fuel economy while maintaining the drivability of the vehicle. Therefore, a possible criterion of predictive control would be the minimization of a quadratic cost function of the control error and the controlled signal, respectively, during the prediction and control horizons. Hence, the quadratic cost function will be [9]:

\[
J = \sum_{i=N_1}^{N_2} \lambda_{yi} [y_r(k+i) - \hat{y}(k+i|k)]^2 + \sum_{j=1}^{n_u} \lambda_{uj} u^2(k + j - 1)
\]

where \(y_r(k+i)\) and \(\hat{y}(k+i|k)\) are the reference and predicted output signal \(i\) steps ahead and \(u(k + j - 1)\) is the controlled signal \(j - 1\) steps ahead. \(N_1\) and \(N_2\) are the first and last points of the prediction horizon and \(n_u\) is the length of the control horizon. Also, \(\lambda_{yi}\) and \(\lambda_{uj}\) are weighting factors of the control error and controlled signal, respectively.

In order to satisfy the control objectives, in this case, \(y_r\) would be equal to \(E_r\) which is drivers required energy and \(\hat{y}\) equal to \(E\) which is the predicted produced energy by the powertrain. Also, towards minimizations of fuel consumption, \(u\) would be equal to the \(P_{eng}\) which is produced power by ICE, the only consumer of the fuel in the powertrain. Therefore, the mentioned cost function can be written as:

\[
J = \sum_{i=N_1}^{N_2} \lambda_{Ei} [E_r(k+i) - \hat{E}(k+i|k)]^2 + \sum_{j=1}^{n_u} \lambda_{Peng} P_{eng}^2(k + j - 1)
\]

Note that the control objectives should be satisfied while considering several constraints. It is desirable to use the battery energy as much as possible to improve fuel economy. However, if the energy recovered by regenerative braking is not enough to sustain the battery charge, this performance can leave the battery completely discharged at the end of the mission [10]. Thus, the integral constraint of the problem is that the state of charge of the battery (SOC) should be really close to the nominal SOC at the end of the mission. It means:

\[
|SOC(t_f) - SOC_{target}| < \epsilon
\]

where \(\epsilon\) is a small quantity which will be defined in the simulation. To avoid complexity in the problem, it was decided to consider this integral constraint as a soft constraint and add the corresponding penalty function quadratically to the cost function:

\[
J = \sum_{i=N_1}^{N_2} \lambda_{SOCi} [SOC_r(k+i) - \hat{SOC}(k+i|k)]^2 + \lambda_{Ei} [E_r(k+i) - \hat{E}(k+i|k)]^2 + \sum_{j=1}^{n_u} \lambda_{Peng} P_{eng}^2(k + j - 1)
\]

Based on a set of trial and error it was decided to take \(N_1 = 1\) and \(N_2 = 10\). Also, \(SOCr\) in all the moments would be equal to \(SOC_{target}\), so it is a constant. In this case, if the cost function gets translated to matrix, \(J\) would be:

\[
J = \begin{bmatrix}
SOC(k+1) - SOC_r & \ldots & SOC(k+10) - SOC_r \\
SOC(k+1) - SOC_r & \vdots & \vdots \\
SOC(k+1) - SOC_r & \vdots & \vdots \\
E(k+1) - E_{ri} & \ldots & E(k+10) - E_{ri} \\
E(k+1) - E_{ri} & \vdots & \vdots \\
E(k+1) - E_{ri} & \vdots & \vdots \\
\end{bmatrix} 
\]

Because of the limitations of the powertrain’s components, there are several local constraints which are summarized as:

\[
SOC_{min} \leq SOC(k+i) \leq SOC_{max} \\
P_{eng,min} \leq P_{eng}(k+j-1) \leq P_{eng,max} \\
P_{bat,min} \leq P_{bat}(k+j-1) \leq P_{bat,max} \\
P_{brk,min} \leq P_{brk}(k+j-1) \leq P_{brk,max}
\]

where \(P_{bat}\) and \(P_{brk}\) are produced power by the battery and the braking system. The parameters written with \(\text{min}\) and \(\text{max}\) subscripts present the minimum and maximum of the corresponding variables.

III. CONTROL-ORIENTED MODEL

Since a model-based approach has been used, it was needed to derive a precise model to calculate accurate prediction of
the output signals. Also, to design a real-time implementable controller, the model should be relatively simple. Towards this end, (7) is considered as the longitudinal dynamics of the vehicle:

\[ F_{\text{trac}} = F_{\text{pwt}} - F_{\text{brk}} \]  

where \( F_{\text{trac}} \) is the tractive force generated by the powertrain and the brake system. \( F_{\text{pwt}} \) is the force produced by the power train and \( F_{\text{brk}} \) is the force produced by the brakes.

To use the concept of power and energy, (7) can be multiplied by the velocity of the vehicle. Also, by utilizing the approximate relation between power and energy and the fact that \( P_{\text{put}} = P_{\text{eng}} + P_{\text{bat}} \) it can be written as:

\[ \Delta E_{\text{trac}} = \Delta t(P_{\text{eng}} + P_{\text{bat}} - P_{\text{brk}}) \]  

(8)

One of the key components of a HEVs’ powertrain is the battery (or in some cases super-capacitors). Discharging or overcharging can be harmful to the battery. To protect the battery from damage, EMS should monitor the state-of-charge (SOC) and keep it in a safe range. The approximate relation between power and energy and the fact that \( P_{\text{put}} = P_{\text{eng}} + P_{\text{bat}} \) is used to write:

\[ E(k + 1) = E(k) + P_{\text{eng}} + P_{\text{bat}} - P_{\text{brk}} \]  

(12)

\[ SOC(k + 1) = SOC(k) - \frac{V_{oc} \sqrt{V_{oc}^2 - 4R_{bat}P_{bat}(k)}}{2R_{bat}Q_{nom}} \]  

(13)

In these equations, state variables, manipulated variables, and outputs are \( X, U, \) and \( Y \) respectively:

\[
X = \begin{bmatrix} E(k) \\ SOC(k) \end{bmatrix} \\
U = \begin{bmatrix} P_{\text{eng}}(k) \\ P_{\text{bat}} \\ P_{\text{brk}} \end{bmatrix} \\
Y = \begin{bmatrix} E(k) \\ SOC(k) \end{bmatrix}
\]

(14)

IV. ENERGY MANAGEMENT SYSTEM

In this section, to design a controller, it is just needed to tune the weighting factors in the cost function to achieve the control objectives and satisfying the constraints while trying to make a compromise between them according to their priorities. To this end, it won’t be very difficult to tune the weighting factors for a specific drive cycle. However, this tuning won’t be very suitable for other drive cycles, since the controller has been tuned to attain an optimal manipulated variable for a certain drive cycle. Moreover, it takes relatively long time to compute the manipulated variables for each reference signal, specially because the state equations and constraints are nonlinear. Thus, it is not possible to implement the controller in real-time. To decrease the computational load in real-time, it was decided to compute the manipulated variables for a vast horizon of reference signals and initial conditions off-line. In this case, \( E_r \) and \( E_1 \) were considered to vary from -50000 J to 50000
J, $SOC_r$ varies from 0.3 to 0.9, and $SOC_1$ varies from 0.2 to 0.9. After a set of trial and error, it got clear that tuning the weighting factors for the whole horizon of reference signals and initial conditions is not practical. Thus, the whole region gets divided to four sub-region based on the amount of $E_r$, $SOC_r$, $E_1$, and $SOC_1$ with respect to each other.

The first sub-region is the area in which $E_1 \leq E_r$ and $SOC_1 \leq SOC_r$. In this case, powertrain has to generate power, however, state-of-charge is lower than its desired amount. So, it would be ideal that the difference between $E_1$ and $E_r$ can be produced only by the engine.

At the second sub-region where $E_1 \leq E_r$ and $SOC_1 \geq SOC_r$, battery's charge is higher than its required amount, so to reduce the fuel consumption, EMS should provide the demanded energy by the battery and use the engine just in the case that the battery won’t be able to generate whole difference between $E_1$ and $E_r$. At the third sub-region power-train doesn’t need to generate power because $E_r \leq E_1$. However, because $SOC_1 \leq SOC_r$, an amount of $P_{brk}$ should be regenerated to charge the battery. So $P_{bat}$ should be negative in this area.

Finally at the last sub-region where $E_r \leq E_1$ and $SOC_1 \geq SOC_r$, there is no need to generate or regenerate power. Thus, $P_{eng}$ and $P_{bat}$ should be equal to zero and $P_{brk}$ should dissipate the whole difference between $E_1$ and $E_r$ by mechanical brake.

After some trial and error, a set of weighting factors which satisfies objectives was obtained. Then, the optimization problem with the resulted weighting factors is solved and the achieved manipulated variables are stored as look-up tables. Thus, in any instant with any reference signals and any initial condition, the controller can interpolate between look-up tables.

---

**Fig. 4** $P_{eng}$ vs $E_1$ and $SOC_1$ while $SOC_r = 0.5$ and $E_r = 0.7$.

**Fig. 5** $P_{bat}$ vs $E_1$ and $SOC_1$ while $SOC_r = 0.5$ and $E_r = 0.7$.

**Fig. 6** $P_{brk}$ vs $E_1$ and $SOC_1$ while $SOC_r = 0.5$ and $E_r = 0.7$. 
TABLE I

VEHICLE CHARACTERISTICS

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective mass</td>
<td>1300 kg</td>
</tr>
<tr>
<td>Frontal area</td>
<td>1.8 m²</td>
</tr>
<tr>
<td>Drag coefficient</td>
<td>0.32</td>
</tr>
<tr>
<td>Rolling resistance</td>
<td>0.013</td>
</tr>
</tbody>
</table>

and compute the manipulated variables in real-time. Although the difference between two adjacent chosen reference signals is relatively high, the computed manipulated variables were expected to be accurate enough because of the nonlinear precise applied equations.

Fig. 4 shows $P_{\text{eng}}$ versus $E_1$ and $SOC_1$. This figure belongs to the situation in which $E_r = 0$ and $SOC_r = 0$. As we expected, $P_{\text{eng}}$ is higher in the first sub-region than the second one. As the Fig. 5 shows, in the second sub-region EMS has tried to provide the demanded energy by the battery. Also, in the 3rd and the 4th sub-region $P_{\text{eng}}$ is equal with zero and $P_{\text{bat}}$ is maximum negative and zero respectively. As shown in the Fig. 6, everything we expected about $P_{\text{brk}}$ has happened. $P_{\text{brk}}$ is equal to zero in the 1st and the 2nd sub-region and its amount is higher in 4th sub-region than 3rd sub-region. It means some of $P_{\text{brk}}$ has regenerated in 3rd sub-region.

V. SIMULATION

For the simulation, the control-oriented model is being used again. To use look-up table in every moment $E_r$ is taken from the driver demand (in this case from drive cycle). $SOC_r$ is a constant and predetermined quantity. Initial conditions have to be measured from the powertrain. All of the information in any instant are required to be fed into the controller so that it can interpolate between look-up tables and calculate the manipulated variables. As $E_r$, we used two drive cycle: UDDS and HWYCOL. However, these drive cycles demonstrate the velocity of the vehicle versus time. To make these drive cycles usable for this controller, by the help of longitudinal dynamics, they got translated to the profiles which show reference energy versus time. In these simulations, the sample time is 1 second.

The initial SOC and target SOC are defined by $SOC_1 = 0.6$ and $SOC_t = 0.7$ respectively. The characteristics of the vehicle are shown in Table I.

Figs. 7 and 8 show the variation of SOC through the drive cycles. Figs. 9 and 10 show how the vehicle can track the driver’s required energy.

Table III summarizes the initial and final SOC, fuel economy, and mean square of error between reference and generated energy. Final SOC is close enough to the target SOC. A little bit of difference between the target and final SOC at the end of a cycle is acceptable and does not affect the vehicle functionality [2]. As it is observable, the mean square of error is very small compared with the order of reference signal which it is a sign of high drivability of the vehicle.

VI. CONCLUSION

This research described an approach to design an energy management system for a charge-sustaining power-split hybrid electric vehicle based on the nonlinear model predictive control. We did not use any knowledge related to the future driving cycle. Thus, our controller is robust regarding the driving condition. In charge-sustaining operation there is a global constraint dictates that the battery state-of-charge can not be deviate largely from its target value. So we’ve added a nonlinear penalty function of the battery state-of-charge deviation from its target value to the cost function. Our control objectives are improving fuel economy while maintaining the

<table>
<thead>
<tr>
<th>Drive cycle</th>
<th>Initial SOC</th>
<th>Final SOC</th>
<th>Fuel economy ($\text{L}/100\text{km}$)</th>
<th>Mean square of error ($J^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDDS</td>
<td>0.6</td>
<td>0.68</td>
<td>3.14</td>
<td>1.4492</td>
</tr>
<tr>
<td>HWYCOL</td>
<td>0.6</td>
<td>0.67</td>
<td>1.03</td>
<td>7.1227</td>
</tr>
</tbody>
</table>
vehicle drivability which are totally achieved according to the simulation section. Moreover, because for computing the manipulated variables, the controller just needs to interpolate in a database, which occupies a small amount of memory and requires limited CPU capability. Therefore, it is totally implementable in real-time.

REFERENCES


