

REAL-TIME UPDATING AND NEAR-OPTIMAL  
ENERGY MANAGEMENT SYSTEM FOR  
MULTI-MODE ELECTRIFIED POWERTRAIN  
WITH REINFORCEMENT LEARNING  
CONTROL

REAL-TIME UPDATING AND NEAR-OPTIMAL ENERGY  
MANAGEMENT SYSTEM FOR MULTI-MODE ELECTRIFIED  
POWERTRAIN WITH REINFORCEMENT LEARNING  
CONTROL

BY

ATRIYA BISWAS, B.Eng., M.Tech.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MECHANICAL ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

© Copyright by Atriya Biswas, December 2020

All Rights Reserved



Doctor of Philosophy (2020)  
(Mechanical Engineering)

McMaster University  
Hamilton, Ontario, Canada

TITLE: Real-time Updating and Near-Optimal Energy Management System for Multi-mode Electrified Powertrain with Reinforcement Learning Control

AUTHOR: Atriya Biswas  
M.Tech., (Mechanical Engineering)  
Indian Institute of Technology, Kharagpur, India

SUPERVISOR: Ali Emadi

NUMBER OF PAGES: xxxviii, 331

# Abstract

Energy management systems (EMSs), implemented in the electronic control unit (ECU) of an actual vehicle with electrified powertrain, is a much simpler version of the theoretically developed EMS. Such simplification is done to accommodate the EMS within the given memory constraint and computational capacity of the ECU. The simplification should ensure reasonable performance compared to theoretical EMS under real-life driving scenarios. The process of simplification must be effective to create a versatile and utilitarian EMS.

The reinforcement learning-based controllers feature profitable characteristics in optimizing the performance of controllable physical systems as they do not mandatorily require a mathematical model of system dynamics (i.e. they are model-free). Quite naturally, it can aspired to testify such prowess of reinforcement learning-based controllers in achieving near-global optimal performance for energy management system (supervisory) of electrified powertrains. Before deployment of any supervisory controller as a mainstream controller, they should be essentially scrutinized through various levels of virtual simulation platforms with an ascending order of physical system emulating-capability. The controller evolves from a mathematical concept to an utilitarian embedded system through a series of these levels where it undergoes gradual

transformation to finally become apposite for a real physical system. Implementation of the control strategy in a Simulink®-based forward simulation model could be the first stage of the aforementioned evolution process. This brief will delineate all the steps required for implementing an reinforcement learning-based supervisory controller in a forward simulation model of a hybrid electric vehicle. A novel framework of loss-minimization based instantaneous optimal strategy is introduced for the energy management system of a multi-mode hybrid electric powertrain in this brief. The loss-minimization strategy is flexible enough to be implemented in any architecture of electrified powertrains. It is mathematically proven that the overall system loss minimization is equivalent to the minimization of fuel consumption.

An online simulation framework is developed in this article to evaluate the performance of a multi-mode electrified powertrain equipped with more than one power source. An electrically variable transmission with two planetary gear-set has been chosen as the centerpiece of the powertrain considering the versatility and future prospects of such transmissions. It is noteworthy to mention that a novel architecture topology selected for this dissertation is engendered through a series of rigorous screening process whose workflow is presented here with brevity. One of the legitimate concern of multi-mode transmission is it's proclivity to contribute discontinuity of power-flow in the downstream of the powertrain. Mode-shift events can be predominantly held responsible for engendering such discontinuity. Advent of dynamic coordinated control as a technique for ameliorating such discontinuity has been substantiated by many scholars in literature. Hence, a system-level coordinated control is employed within the energy management system which governs the mode schedule of the multi-mode powertrain in real-time simulation.

*I dedicate this dissertation to my Maa, Baba, and Didi, who have inspired my direction.*

# Acknowledgements

This research was undertaken, in part, thanks to funding from the Canada Excellence Research Chairs Program and Natural Sciences and Engineering Research Council of Canada (NSERC). I would like to thank Dr. Ali Emadi for giving me the chance to be a member of his research group and the opportunity to do my Ph.D studies under his supervision. I feel blessed to get Dr. Emadi's encouragement in improving the quality and plan of my research path.

The long meandering road of a doctorate is never possible without the support of my loving family. There is no words to express my gratitude towards my parent and my sister for their immense love and support. Whatever today I have achieved, nothing would have been possible without my mother's relentless encouragement and passion for knowledge.

I would like to extend my profound appreciation to my committee members; Dr. Fengjun Yan and Dr. Ryan Ahmed for their leadership during my research through their encouragement and the constructive feedback.

I would like to thank Dr. Jennifer Bauman for her constructive feedback in the comprehensive examination and during her course. I also wish to show my gratitude to Dr. Berker Bilgin for helping in my research in the early days of graduate study. I would like to thank Dr. Phillip Kollmeyer for meticulously reviewing and helping

to publish my first academic paper. I would like thank Dr. Ran Gu for academic support.

I wish to show my gratitude to Teresa Janes, Theresa Mitchell, Paul Nguyen and Dan Manolescu for managing and coordinating the research needs. I cannot finish without acknowledging Nicole Mclean, Leslie Kocsis, Lily Sazz-Fayter for being wonderful graduate administrator team of mechanical engineering department. I would like to thank Dr. Mukesh K. Jain, Dr. Chan Y. Ching, and John Colenbrander for being my TA supervisor and helping me in sharpening my student management skill.

In addition, I wish to express my deepest gratitude to Pier Giuseppe Anselma for being the research partner throughout the critical stages of my graduate study. I also would like to thank Joel Roeleveld and Omkar Rane for their help with the concepts of powertrain electrification. It would be unfair if I don't acknowledge the help from Mathworks (Maharshi Patel, Russell Graves) for identifying the bug in their toolbox and helping me in learning reinforcement learning toolbox. Furthermore, I would like to thank my close friends in the research group: Dr. Peter Azer, Sumedh Dhale, Carin Yeghiazarian, Lucas Bruck, Yue Wang, Aashit Rathore, Saeed Amirfarhangi Bonab, Jeremy Lampert, Alex Wang, Dr. Aniket Anand. I also want to thank my colleagues in the lab: Reihaneh Ostadian Bidgoli, Adam Lampert, Jianbin Liang, Fabricio Machado, Zekun Xia, Diego Fernando Valencia, Josh Taylor, Mohamed Ibrahim, Dr. Romina Rodriguez, Maryam Alizadeh, John Ramoul, Dr. Carlos Vidal, Parisa M Shamsabadi, Dr. Shamsuddeen Nalakath, Dr. Ephrem Chemali, Dr. Alan Callegaro, Alex Allca-Pekarovic, Jigar Mistry, and Jack Gillies.

I want to thank my friends in Hamilton: Dipankar Saha, Krishna Jangid, Sudarshan Sharma, Ramis Arbi, Hosnee Mobarak, and Chirayu Chokshi.

Some people touch our lives in a way that cannot be expressed, making an immortal impact, both in life and death. I acknowledge the influence of Iman and Mehdi by commemorating the time shared together in MARC 218. I wish they would have graduated along with me, just like it was planned.

Finally and above all, I would like to thank God Almighty for giving me the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it satisfactorily. Without his blessings, this achievement would not have been possible.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>vi</b>
<b>Abbreviations and Symbols</b>	<b>xxix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Reinforcement Learning-based Control for Energy Management System	2
1.2 Motivation . . . . .	9
1.3 Objectives and Contributions . . . . .	11
1.4 Summary of Thesis . . . . .	12
<b>2 Electrified Powertrain: Fundamentals and Classification</b>	<b>16</b>
2.1 Introduction . . . . .	17
2.2 Requirement of Electrified Powertrains . . . . .	17
2.3 Classification Electrified Powertrain Architectures . . . . .	19
2.4 Summary . . . . .	35
<b>3 Energy Management Systems: Fundamentals, Classification, and Utilitarian Aspect</b>	<b>37</b>



3.1	Introduction . . . . .	38
3.2	Energy Management System based on Powertrain Architecture . . . . .	39
3.3	Fundamentals, Evolution, and classification of energy management systems . . . . .	41
3.4	Premeditated Energy Management Systems . . . . .	48
3.5	Casual Energy Management Systems . . . . .	60
3.6	Blended Energy Management Systems . . . . .	63
3.7	Development of Utilitarian Energy Management System . . . . .	66
3.8	Future Trends: Online Update of Energy Management Systems . . . . .	72
3.9	Summary . . . . .	75
<b>4</b>	<b>Articulation of an Instantaneous Optimization based Control Strategy</b>	<b>79</b>
4.1	Introduction . . . . .	80
4.2	Equivalent Consumption Minimization Strategy: A Casual Incarnation of Pontryagin’s Minimization Problem . . . . .	81
4.3	Implementation of Optimal ECMS . . . . .	85
4.4	Summary . . . . .	108
<b>5</b>	<b>In Pursuit of a Novel Strategy: Overall System Loss Minimization Strategy instead of Engine’s Efficiency Maximization</b>	<b>110</b>
5.1	Introduction . . . . .	111
5.2	Fundamentals of Loss Minimization Strategy . . . . .	111
5.3	Revised Cost Structure . . . . .	115
5.4	Simulation and Results . . . . .	120

5.5	Summary . . . . .	131
<b>6</b>	<b>Introduction of Reinforcement Learning in Solving the Energy Management Problem of Hybrid Electric Vehicle</b>	<b>138</b>
6.1	Introduction . . . . .	139
6.2	Fundamental of RL based Control Strategy . . . . .	140
6.3	Application of RL Agent in Real-Time Control for EMSs . . . . .	150
6.4	Performance of Energy Management System with Q-Learning based RL Agent . . . . .	163
6.5	summary . . . . .	166
<b>7</b>	<b>Application of Asynchronous Actor-Critic Agent in Real-time Energy Management of a Multi-mode Electrified Powertrain</b>	<b>170</b>
7.1	Introduction . . . . .	171
7.2	Inspiration for Online Simulation of Asynchronous Advantage Actor-Critic Agent . . . . .	173
7.3	Contribution . . . . .	176
7.4	Requirement and Fundamentals of Functional Approximation . . . . .	177
7.5	Training of Deep Reinforcement Learning Agents . . . . .	182
7.6	Multi-mode Powertrain Modeling for RL Application . . . . .	188
7.7	Online Implementation of Asynchronous Advantage Actor-Critic Agent as EMS Controller . . . . .	191
7.8	Results and Discussion . . . . .	202
7.9	summary . . . . .	221

<b>8</b>	<b>Advanced Modeling of Multi-mode Electrified Powertrain: Requirement of a Dynamic Coordinated Control to Manage the Mode-shift Events Efficiently</b>	<b>227</b>
8.1	Introduction . . . . .	228
8.2	Motivation . . . . .	229
8.3	Automated Topology Selection Tool . . . . .	233
8.4	Powertrain Modeling . . . . .	236
8.5	Clutch-based Coordinated Control during Mode-shifts . . . . .	243
8.6	Instantaneous Optimization-based Mode Scheduling Control for Multi-mode ePT . . . . .	256
8.7	Simulation Results and Discussion . . . . .	259
8.8	Summary . . . . .	272
<b>9</b>	<b>Conclusions, Future Work and Publications</b>	<b>278</b>
9.1	Conclusions . . . . .	279
9.2	Future Work . . . . .	281
9.3	Publications . . . . .	285
	<b>Appendices</b>	<b>288</b>
<b>A</b>	<b>Efficiency and fuel contour maps of a few ICE models, SOC-OCV and efficiency contour map of the battery used in this dissertation</b>	<b>289</b>
<b>B</b>	<b>Detailed Pseudo-code for Q-learning algorithm used for offline simulation in this dissertation</b>	<b>297</b>
	<b>References</b>	<b>301</b>

# List of Figures

1.1	HEV sales trend in U.S. from 1999 to 2017 [5]. . . . .	3
1.2	Sales trend of drivetrain architectures in U.S. from 2011 to 2017 [5]. . . . .	4
1.3	Curse of dimensionality . . . . .	9
1.4	Sluggishness of convergence for traditional RL algorithms with discrete-time domain. . . . .	10
1.5	Unpredictable parameters such as traffic situations and driving routes to the destination during real-world driving scenario. . . . .	11
2.1	Automotive sales statistics in last one and a half decades [1]. . . . .	18
2.2	U.S. GHG emissions by economic sector, 2017 and GHG statistics for transportation sector [29]. . . . .	19
2.3	Variation of Canada’s GHG emissions since 1990 to 2018 in different economic sectors and GHG statistics for transportation sector [30]. . . . .	20
2.4	Hierarchical classification of architectures of electrified powertrains. . . . .	20
2.5	Architecture of a micro hybrid electric powertrain with less than 5% electrification factor. . . . .	21
2.6	Architecture of a mild hybrid electric powertrain with (10%-15%) electrification factor. . . . .	22

2.7	Different configurations of a P1 hybrid electric powertrain architecture with (20%-40%) electrification factor. . . . .	23
2.8	Architecture of a P4 hybrid electric powertrain with (20%-40%) electrification factor. . . . .	24
2.9	Architecture of a series hybrid electric powertrain with (40%-50%) electrification factor. . . . .	25
2.10	Architecture of a power-split hybrid electric powertrain with (20%-40%) electrification factor. . . . .	26
2.11	Schematic diagram of first “Two-mode hybrid system” with single PG-set. . . . .	28
2.12	Schematic diagram of advanced “Two-mode hybrid system” with two PG-sets. . . . .	29
2.13	Schematic diagram of Single-Input Electrically Variable Transmission with two PG-sets and the modes facilitated by it. . . . .	30
2.14	HEV sales trend in U.S. from 1999 to 2017. . . . .	34
3.1	Representative optimum ICE operating zones with respect to different criteria for a typical ICE map. . . . .	43
3.2	Tentative evolution of EMSs since 1993 till 2018. . . . .	45
3.3	Trend of popularity of major control strategies used in EMSs . . . . .	47
3.4	Primary components of premeditated EMS . . . . .	49
3.5	Example of deterministic rule-based control policy in premeditated EMS	50
3.6	Example of fuzzy rule-based control policy in premeditated EMS . . .	50
3.7	Typical look-up table based premeditated EMS . . . . .	51
3.8	Basic articulation steps for premeditated EMS with post-processing .	53

3.9	Cyclical repetition of policy evaluation and improvement . . . . .	58
3.10	Example of control policy obtained from multi-parametric programming	66
3.11	Chronological steps of EMS development . . . . .	67
3.12	Schematic of hardware-in-the-loop simulation set-up . . . . .	70
3.13	Concept of online update of control strategy with onboard computation	73
3.14	Concept of online update of control strategy for energy management system . . . . .	74
3.15	Concept of online update of control strategy for energy management system . . . . .	77
4.1	Overview of the simulation framework used for ECMS implementation	86
4.2	Schematic diagram of the driver block used in a model-in-the-loop sim- ulation for vehicles . . . . .	87
4.3	Schematic diagram of the three modes facilitated by the multi-mode powertrain chosen for this study . . . . .	88
4.4	Numbering of nodes of a two-planetary gear set system with an example	91
4.5	Methodology for obtaining constraints matrix, with an example. . . .	92
4.6	Constraints matrix for the powertrain architecture of Chrysler Pacifica with explanation. . . . .	93
4.7	Transition “M” matrix for the powertrain architecture of Chrysler Paci- fica for the HEV and single-motor all-electric modes. . . . .	94
4.8	Transition “P” matrix for the powertrain architecture of Chrysler Paci- fica for the HEV and single-motor all-electric modes. . . . .	95
4.9	Transition “M” matrix for the powertrain architecture of Chrysler Paci- fica for the two-motor all-electric mode. . . . .	96

4.10	Transition “P” matrix for the powertrain architecture of Chrysler Pacifica for the two-motor all-electric mode. . . . .	97
4.11	Mass flow rate of fuel consumption, wide open throttle (WOT) torque, and friction torque. . . . .	99
4.12	Efficiency contour plots of traction motor at different voltages. . . . .	101
4.13	Efficiency contour plots of generator at different voltages. . . . .	102
4.14	Overview of the ECMS controller block in a model-in-the-loop simulation with a multi-mode electrified powertrain. . . . .	103
5.1	Power flow in multi-mode power-split electrified powertrain . . . . .	113
5.2	Comparison between battery SOC trajectories when inertia-based plant model is used instead of steady-state plant . . . . .	122
5.3	Increase of cumulative fuel consumption when an inertia-based plant model is used instead of a steady-state plant for a same drive cycle . . . . .	123
5.4	Comparison of IC Engine speed response between steady-state and inertia-based plant model . . . . .	124
5.5	Vehicle speed, fuel consumption, and SOC trajectories for UDDS . . . . .	125
5.6	Comparison of powertrain operating modes under LMS and ECMS for UDDS . . . . .	126
5.7	Comparison of ICE speed under LMS and ECMS for UDDS . . . . .	127
5.8	Comparison of ICE torque under LMS and ECMS for UDDS . . . . .	128
5.9	Vehicle speed, fuel consumption, and SOC trajectories for WLTC . . . . .	129
5.10	Comparison of powertrain modes under LMS and ECMS for WLTC . . . . .	130
5.11	Comparison of ICE speed under LMS and ECMS for WLTC . . . . .	130
5.12	Comparison of ICE torque under LMS and ECMS for WLTC . . . . .	131

5.13	Vehicle speed, fuel consumption, and SOC trajectories for HWFET . . . . .	132
5.14	Comparison of powertrain mode under LMS and ECMS for HWFET . . . . .	133
5.15	Comparison of ICE speed under LMS and ECMS for HWFET . . . . .	133
5.16	Comparison of ICE torque under LMS and ECMS for HWFET . . . . .	134
5.17	Time-series of mode-shift events throughout WLTC under loss minimization strategy without drivability constraint . . . . .	134
5.18	Time-series of mode-shift events throughout WLTC under loss minimization strategy with drivability constraint . . . . .	134
5.19	Comparison of fuel consumption throughout WLTC under loss minimization strategy with and without drivability constraint . . . . .	135
5.20	Torque and speed responses of high-fidelity ICE w.r.t controller's command . . . . .	136
5.21	Torque and speed dynamics of ICE during cranking and HEV deactivation . . . . .	137
6.1	Conceptual diagram of sequential decision making, and relationship among state, action, and reward in discrete time-step. . . . .	141
6.2	Schematic representation of sequential decision problem. . . . .	147
6.3	Agent-environment interaction for Q-learning. . . . .	150
6.4	Schematic of Q-value function update with continuous policy improvement. . . . .	151
6.5	Concept of periodical update of control policy with on-board ECU or computer. . . . .	156
6.6	Concept of periodic update of control policy with off-board high-performance computer. . . . .	157



6.7	a) Schematic configuration of Toyota Prius 2010 powertrain, b) Lever diagram of the hybrid-electric mode, and c) Lever diagram of the single-motor pure electric mode. . . . .	158
6.8	Tentative map of brake specific fuel consumption of the Atkinson cycle-based ICE used in Toyota Prius III generation. . . . .	159
6.9	Tentative maps of the battery pack's Open circuit voltage and the battery pack's internal resistance are expressed as a mathematical function of battery SOC, used in Toyota Prius. . . . .	160
6.10	Tentative efficiency map of the traction motor used in Toyota Prius's electrified powertrain. . . . .	161
6.11	Tentative efficiency map of the generator used in Toyota Prius's electrified powertrain. . . . .	162
6.12	Algorithm used to calculate immediate cost for every combination of state and action. . . . .	163
6.13	Part.1: RL algorithm for simultaneous policy iteration and Q-value iteration. . . . .	165
6.14	Part.2: RL algorithm for simultaneous policy iteration and Q-value iteration. . . . .	166
6.15	Result of Q-learning: Ability of full cycle completion with charge sustenance. . . . .	167
6.16	Trend of SOC profiles obtained from different simulations of a Q-learning-based RL agent trained with 800 iterations. . . . .	167
6.17	Time-series of engine speed obtained from a Q-learning-based RL agent trained with 800 iterations. . . . .	168

6.18	Time-series of engine torque obtained from a Q-learning-based RL agent trained with 800 iterations. . . . .	168
6.19	Comparison of performance between DP and a Q-learning-based RL agent trained with 800 iterations. . . . .	169
7.1	Main contribution of this dissertation: Application of asynchronous actor-critic in yielding near-global optimal control in real-world driving scenario. . . . .	172
7.2	Schematic diagram of forward-approach and backward-approach simulation model. . . . .	175
7.3	Schematic diagram of actor-critic based agent with nonlinear function approximation. . . . .	181
7.4	Evolution of ANN-based function approximator as RL agent. . . . .	185
7.5	Schematic diagram of three distinct modes facilitated by the eVT transaxle. . . . .	189
7.6	Schematic diagram of multi-mode eVT transaxle. . . . .	190
7.7	Conflation of RL and ECMS in online implementation of multi-mode EMS. . . . .	197
7.8	Deep neural network architecture for actor and critic network. . . . .	199
7.9	Sequence of drive cycles for training schedule. . . . .	202
7.10	Range of vehicle's power demand over 58 real-world drive cycles. . . . .	204
7.11	Random drive cycle generation with Markov chain model. . . . .	206
7.12	Acquired and randomly generated drive cycles used for training. . . . .	207
7.13	Training history of A3C agent-based EMS controller through 1000 episodes. . . . .	208

7.14	a) Randomly generated drive cycle, which is used to testify the A3C agent, plotted with right axis. Battery SOC variation throughout the drive cycle is plotted with left axis. b) Time-series of powertrain mode and fuel consumption are plotted with left axis and right axis respectively.	209
7.15	Time-series of $\omega_{ice}$ , $\omega_{gen}$ , and $\omega_{mot}$ throughout the real-world drive cycle.	210
7.16	Time-series of $\tau_{ice}$ , $\tau_{gen}$ , and $\tau_{mot}$ throughout the real-world drive cycle.	211
7.17	Comparison of powertrain mode profiles between ECMS and A3C agent when applied on real-world drive cycle. . . . .	212
7.18	Comparison of battery SOC profiles between ECMS and A3C agent when applied on real-world drive cycle. . . . .	212
7.19	Comparison of fuel consumption between ECMS and A3C agent when applied on real-world drive cycle. . . . .	213
7.20	Time-series of powertrain mode obtained for UDDS with A3C agent as an EMS control strategy. . . . .	213
7.21	Time-series of $\tau_{ice}$ , $\tau_{gen}$ , and $\tau_{mot}$ obtained for UDDS with A3C agent as an EMS control strategy. . . . .	214
7.22	Time-series of $\omega_{ice}$ , $\omega_{gen}$ , and $\omega_{mot}$ obtained for UDDS with A3C agent as an EMS control strategy. . . . .	215
7.23	Time-series of battery SOC variation throughout UDDS with A3C agent as an EMS control strategy. . . . .	215
7.24	Effect of training iteration numbers on the agent's convergence to near-global optimality. . . . .	218

7.25	Effect of training iteration numbers on the agent’s performance (drivability, charge-sustenance, and fuel-consumption) compared to optimal ECMS. . . . .	219
7.26	Effect of training iteration numbers on the agent’s fuel consumption performance compared to optimal ECMS for real-world drive cycle. . . . .	223
7.27	Effect of training iteration numbers on the agent’s fuel consumption performance compared to optimal ECMS for UDDS. . . . .	224
7.28	Effect of training iteration numbers on the agent’s fuel consumption performance compared to optimal ECMS for WLTP. . . . .	224
7.29	Effect of training iteration numbers on the agent’s charge sustenance capability. A real-world drive cycle is chosen here. . . . .	225
7.30	Effect of training iteration numbers on the agent’s charge sustenance capability. UDDS is chosen here. . . . .	226
7.31	Effect of training iteration numbers on the agent’s charge sustenance capability. WLTP is chosen here. . . . .	226
8.1	Example of a specific powertrain topology derived from a specific powertrain configuration. . . . .	229
8.2	Automatic and exhaustive topology candidate screening process. . . . .	234
8.3	Outcome of a rapid assessment of fuel economy capability of more than 100,000 topology candidates. . . . .	236
8.4	Mode-shifts among the four feasible modes. . . . .	237
8.5	Systematic method of obtaining dynamic relationship among components and nodes of an ePT with multiple PG-sets. . . . .	239

8.6	Different stages of clutch-to-clutch shift from mode#1 to mode#3 in a case when $\omega_{ice}$ is greater than $\omega_{out}$ . . . . .	246
8.7	Three cardinal stages of torque transfer during mode-shift from mode#1 to mode#3. . . . .	247
8.8	Three cardinal stages of torque transfer during mode-shift from mode#3 to mode#1. . . . .	249
8.9	Kinematic and dynamic relationship among nodes of PG-sets during 2-to-4 mode-shift. . . . .	250
8.10	Approximate clutch torque variation of CL#1 and CL#2 during 2-to-4 mode-shift. . . . .	251
8.11	Kinematic and dynamic relationship among nodes of PG-sets during 4-to-2 mode-shift. . . . .	252
8.12	Approximate clutch torque variation of CL#1 and CL#2 during 4-to-2 mode-shift. . . . .	253
8.13	internal combustion engine (ICE) characteristics curves required for modeling the ICE. (a) Mean-value wide open throttle torque as a function of ICE speed. (b) Mass-flow rate of fuel consumption as a function of ICE speed and torque. (c) Mean-value ICE resistance torque as a function of ICE speed. . . . .	256
8.14	Implementation of coordinated control in a multi-mode ePT where mode scheduling is governed by ECMS . . . . .	259

8.15	Trajectories of major components' energy and major losses in 1st All-electric mode (Mode#3) under WLTC drive cycle. The total energy of HVB and the total losses incurred by the HVB, generator, motor, transmission, and road load at the end of the drive cycle are furnished here to support the energy balance presented in Tab.8.5. . . . . .	260
8.16	Trajectories of major components' energy and major losses in 1st Hybrid-electric mode (Mode#1) under WLTC drive cycle. The total energy supplied by the ICE and HVB and the total losses incurred by the HVB, generator, motor, transmission, and road load at the end of the drive cycle are furnished here to support the energy balance presented in Tab.8.5. . . . . .	261
8.17	Comparison of fuel consumption and battery SOC trajectories between inertia-based and steady-state electrified powertrain (ePT) models simulated with same EMS strategy. . . . . .	262
8.18	Timeseries of vehicle speed, $\omega_{ice}$ , $\omega_{mot}$ , $\omega_{gen}$ as a result of ECMS without any coordinated control. . . . . .	263
8.19	Timeseries of vehicle speed, $\omega_{ice}$ , $\omega_{mot}$ , $\omega_{gen}$ as a result of ECMS assisted with partial coordinated control, i.e., during EV-to-HEV and HEV-to-EV mode shifts. . . . . .	264
8.20	Timeseries of vehicle speed, $\omega_{ice}$ , $\omega_{mot}$ , $\omega_{gen}$ as a result of ECMS assisted with full coordinated control, i.e., during EV-to-HEV, HEV-to-EV, HEV-to-HEV, and EV-to-EV mode shifts. . . . . .	265

8.21	Comparison of SOC and fuel consumption trajectories among three aforementioned cases, i.e., ECMS without coordinated control, ECMS with partial coordinated control, and ECMS with full coordinated con- trol. . . . .	266
8.22	Time-series of optimal powertrain operating mode yielded by ECMS assisted by full coordinated control for an entire real-world drive cycle.	267
8.23	Comparison between time-series of $\tau_{ice_{optim.}}$ yielded by “full coordinated control assisted ECMS” and ECMS only for an entire real-world drive cycle. . . . .	268
8.24	Comparison between time-series of $\tau_{gen_{optim.}}$ yielded by “full coordi- nated control assisted ECMS” and ECMS only for an entire real-world drive cycle. . . . .	271
8.25	Battery SOC trajectory yielded with ECMS assisted by full coordi- nated control for an entire real-world drive cycle. . . . .	271
8.26	Comparison between time-series of $\tau_{mot_{optim.}}$ yielded by “full coordi- nated control assisted ECMS” and ECMS only for an entire real-world drive cycle. . . . .	272
8.27	Correlation between ICE speeds yielded from vehicle plant catered with “full coordinated control assisted ECMS” and the speed requested by EMS. . . . .	272
8.28	Correlation between generator speeds yielded from vehicle plant catered with “full coordinated control assisted ECMS” and the speed requested by EMS. . . . .	273

8.29	Correlation between traction motor speeds yielded from vehicle plant catered with “full coordinated control assisted ECMS” and the speed requested by EMS. . . . .	273
A.1	Fuel map of a 2 Liter spark-ignition ICE used in the multi-mode electrified powertrain’s simulation in this dissertation. The ICE’s data is obtained from U.S. Environmental Protection Agency’s (EPA’s) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. This ICE data have been used for the articulation of reinforcement learning agent’s simulation. . . . .	290
A.2	Efficiency map of a 2 Liter spark-ignition ICE used in the multi-mode electrified powertrain’s simulation in this dissertation. The ICE’s data is obtained from U.S. Environmental Protection Agency’s (EPA’s) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. This ICE data have been used for the articulation of reinforcement learning agent’s simulation. . . . .	291
A.3	Some extra Fuel map of another 2 Liter spark-ignition ICE. The ICE’s data is obtained from U.S. Environmental Protection Agency’s (EPA’s) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. . . . .	292
A.4	Some extra efficiency map of another 2 Liter spark-ignition ICE. The ICE’s data is obtained from U.S. Environmental Protection Agency’s (EPA’s) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. . . . .	293



A.5	Some extra Fuel map of another 2.7 Liter spark-ignition ICE. The ICE's data is obtained from U.S. Environmental Protection Agency's (EPA's) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. . . . .	294
A.6	Some extra efficiency map of another 2.7 Liter spark-ignition ICE. The ICE's data is obtained from U.S. Environmental Protection Agency's (EPA's) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. . . . .	295
A.7	SOC-OCV relationship of a Panasonic battery cell used in the multi-mode electrified vehicle used for online simulation. . . . .	296
A.8	Efficiency of a Panasonic battery cell used in the multi-mode electrified vehicle used for online simulation. . . . .	296

# List of Tables

2.1	Clutch and brake status in different modes in the predecessor “Two-mode Hybrid system” with Single PG-set . . . . .	36
2.2	Clutch and brake status in different modes in the successor “Two-mode Hybrid system” with double PG-sets . . . . .	36
3.1	Comparison of control strategies used in EMSs . . . . .	78
4.1	Specification of the vehicle and cardinal components . . . . .	109
4.2	Tuning factor for mode-shift penalty term . . . . .	109
5.1	Comparison of simulation results . . . . .	135
6.1	Vehicle specification of Toyota Prius in brief . . . . .	159
6.2	Discretized values of Vehicle’s Power Demand, Battery SOC, ICE Torque, and ICE Speed . . . . .	164
7.1	Upper and lower bounds of state variables . . . . .	215
7.2	Feasible combinations Control variables of different modes . . . . .	216
7.3	Samples of feasible and infeasible combinations of 4-D action variables	217
7.4	Hyper-parameters for implementing asynchronous actor-critic agent in Simulink® . . . . .	217
7.5	Definition of Transition Probability Matrix . . . . .	218

7.6	Effect of training iteration numbers on the agent’s performance (drivability, charge-sustenance, and fuel-consumption) compared to optimal ECMS . . . . .	220
8.1	Specification of the vehicle and cardinal components . . . . .	275
8.2	Description of modes facilitated by the new electrified powertrain topology . . . . .	275
8.3	Required clutch engagement and disengagement for different mode-shifts (infeasible mode-shifts are shown in red color) . . . . .	276
8.4	Tuning factor for mode-shift penalty term . . . . .	276
8.5	Energy balance as measure of validating plant model . . . . .	277

# Abbreviations and Symbols

## Abbreviations

<b>IOC</b>	Instantaneous optimal control
<b>AI</b>	Artificial intelligence
<b>EMS</b>	Energy management systems
<b>ePT</b>	Electrified powertrain
<b>ECMS</b>	Equivalent Consumption Minimization Strategy
<b>GHG</b>	Greenhouse gases
<b>EV</b>	Electric vehicle
<b>HEV</b>	Hybrid electric vehicle
<b>PHEV</b>	Plug-in-hybrid electric vehicle
<b>EMG</b>	Electric motor/generator
<b>EESS</b>	Electric energy storage system

<b>GOC</b>	Global optimal control
<b>PMP</b>	Pontryagin's minimum principle
<b>EF</b>	Equivalence factor
<b>DP</b>	Dynamic Programming
<b>A-ECMS</b>	Adaptive ECMS
<b>SDP</b>	Stochastic dynamic programming
<b>PID</b>	proportional-integral-derivative
<b>ANN</b>	Artificial neural network
<b>SOC</b>	State-of-charge
<b>OCV</b>	Open circuit voltage
<b>IR</b>	Internal resistance
<b>LMS</b>	Loss minimization strategy
<b>e-CVT</b>	Electrically continuous variable transmission
<b>ICE</b>	Internal combustion engine
<b>WOT</b>	Wide open throttle
<b>SI</b>	Spark-ignition
<b>MFR</b>	Mass flow rate
<b>FAER</b>	Fuel-air equivalent ratio

<b>EECU</b>	Engine electronic control unit
<b>FWD</b>	Forward wheel drive
<b>PG</b>	Planetary gear
<b>OWC</b>	One-way clutch
<b>THS</b>	Toyota hybrid system
<b>TMR</b>	Torque multiplication ratio
<b>MMI</b>	Mass moment of inertia
<b>SIL</b>	Software-in-the-loop
<b>HIL</b>	Hardware-in-the-loop
<b>MIL</b>	Model-in-the-loop
<b>FSM</b>	Forward-approach simulation model
<b>BSM</b>	Backward-approach simulation model
<b>DOF</b>	Degree-of-freedom
<b>RL</b>	Reinforcement learning
<b>LUT</b>	Look-up table
<b>DNN</b>	Deep neural network
<b>TPM</b>	Transition probability matrix
<b>TD</b>	Temporal difference

<b>AI</b>	Artificial intelligence
<b>SGD</b>	Stochastic gradient-descent
<b>GPU</b>	Graphical processing unit
<b>A2C</b>	Asynchronous actor-critic
<b>HVB</b>	High voltage battery
<b>UDDS</b>	Urban dynamometer driving schedule
<b>FTP</b>	Federal test procedure
<b>WLTC</b>	Worldwide harmonized light vehicles test cycle
<b>HWFET</b>	Highway fuel economy test cycle
<b>MDP</b>	Markov decision problem
<b>GA</b>	Genetic algorithm
<b>MCM</b>	Markov chain model
<b>RGDC</b>	Randomly generated drive cycle
<b>ATST</b>	Automated topology selection tool
<b>FBD</b>	Free body diagram
<b>DCCS</b>	Dynamic coordinated control strategy
<b>EPA</b>	Environmental protection agency
<b>OEM</b>	Original equipment manufacturer

<b>GM</b>	General Motors
<b>PEARS</b>	Power-weighted efficiency analysis for rapid sizing
<b>SERCA</b>	Slope-weighted energy-based rapid control analysis
<b>SVM</b>	Support vector machine
<b>VCU</b>	Vehicle control unit
<b>MPC</b>	Model predictive control
<b>QP</b>	Quadratic Programming
<b>PSD</b>	Power-split device
<b>APU</b>	Auxiliary power unit
<b>PSO</b>	Particle swarm optimization
<b>BSG</b>	Belt-driven starter-generator
<b>ISG</b>	Integrated starter-generator
<b>ECU</b>	Electronic control unit
<b>CAN</b>	Controller area network
<b>EVT</b>	Electrically variable transmission
<b>BSFC</b>	Brake specific fuel consumption
<b>OSP</b>	Online simulation platform (Optimization)



## Symbols

$\tau_{ice}$	Torque of internal combustion engine
$\tau_{EMG}$	Torque of electric motor/generator
$\tau_{out}$	Torque at the transmission output
$\tau_{wheel}$	Load torque at the wheel
$\omega_{ice}$	Rotational speed of internal combustion engine
$\omega_{wheel}$	Rotational speed of wheel
$\xi$	Torque-split ratio between ICE and electric machine in parallel HEV
$Pow_{batt}, Pow_{EMG}$	Power of battery, Power of electric machines
$\eta_{EMG}$	Efficiency of electric machines
$\tau_{sun}$	Torque at the sun gear of a PG-set
$\tau_{gen}$	Torque at the generator (the smaller EMG in a powertrain comprising of two EMGs)
$\tau_{mot}$	Torque at the traction motor (the bigger EMG in a powertrain comprising of two EMGs)
$J_{mot}$	MMI of traction motor.
$\theta_{mot}$	Angle of rotation of traction motor

$\tau_{carrier}$	Torque at the carrier of a PG-set
$\tau_{ring}$	Torque at the ring gear of a PG-set
$J_{ice}$	MMI of internal combustion engine
$\theta_{ice}$	Crank angle of internal combustion engine
$N_{ring}$	Number of teeth on the ring gear
$N_{sun}$	Number of teeth on the sun gear
$\beta$	Ratio between $N_{ring}$ and $N_{sun}$ , $\beta_1, \beta_2, \dots, \beta_n$ will represent the ratio for 1 <sup>st</sup> , 2 <sup>nd</sup> , .. $n^{th}$ PG-sets, respectively.
$J_{out}$	MMI of transmission output
$J_{out_{eq}}$	MMI of transmission output when the MMI of the entire vehicle-load is also calculated at the transmission output.
$\theta_{out}$	Angle of rotation of the transmission output shaft
$J_{mot_{eq}}$	Equivalent MMI of the traction motor when it is connected with any of the PG-set components
$J_{gen_{eq}}$	Equivalent MMI of the generator when it is connected with any of the PG-set components
$J_{ice_{eq}}$	Equivalent MMI of the internal combustion engine when it is connected with any of the PG-set components
$\vartheta$	Road grade in radian.

$\varphi_1, \varphi_2$	Two road load coefficients
$C_d$	Aero-drag coefficient
$A_{front}$	Frontal area of any vehicle $0.86 \times height_{vehicle} \times width_{vehicle}$
$r_{wheel}$	Wheel radius in meter.
$\tau_{diff. loss}$	Torque-loss in the differential
$\tau_{mesh loss}$	Torque-loss in the transmission due to meshing of gears
$\tau_{spin. loss}$	Torque-loss in the transmission due to spinning of gears
$\dot{m}_{fuel}$	Mass-flow rate of combustible fuel for internal combustion engine
$x$	state variable for any dynamic system
$\psi(x(t_f))$	Terminal cost based on state variable's terminal status. It represents a penalty due to deviation from end-SOC constraint in EMS.
$P_{load}$	Load on the powertrain in terms of power demand
$u(t)$	A vector comprising control action variables
$\lambda(t)$	Co-state in Pontryagin's Minimum Principle
$H(x, u, t, \lambda(t))$	Time-variant Hamiltonian function in Pontryagin's Minimum Principle. It is a function of state, control, and co-state variables.
$\pi$	Policy of reinforcement learning (RL) agent
$\mathbb{E}$	This symbol is used to express something's expected value.

$v(s)$	State value function for an RL agent. $s$ is representing state of the RL agent.
$q_{s,a}$	Action value function for an RL agent. $a$ is representing control actions of the agent.
$\gamma$	Discount factor of the RL agent
$\alpha$	Learning rate of the RL agent
$v_*(s)$	Minimum value function of state $s$ considering a minimization problem
$q_*(s, a)$	Minimum action value function for state $s$ and action $a$ considering a minimization problem
$\lambda^*$	Optimal value of co-state vector
$\{\tau_0\}$	Generalized torque vector for a multi-mode ePT
$\{\dot{\Omega}_0\}$	Generalized angular acceleration vector for a multi-mode ePT
$\{\tau\}_{HEV/EV1}$	Torque vector as a part of the powertrain dynamics for single-motor mode and hybrid-electric mode in Chrysler Pacifica's Si-EVT
$\{A\}_{HEV/EV1}$	Relationship between torque vector and angular acceleration vector, for single-motor mode and hybrid-electric mode in Chrysler Pacifica's Si-EVT, expressed in matrix form
$\{\tau\}_{EV2}$	Torque vector as a part of the powertrain dynamics for dual-motor mode in Chrysler Pacifica's Si-EVT

$\{A\}_{EV2}$	Relationship between torque vector and angular acceleration vector, for dual-motor mode in Chrysler Pacifica's Si-EVT, expressed in matrix form
$\tau_{frice}$	Friction torque of the internal combustion engine

# Chapter 1

## Introduction

## 1.1 Reinforcement Learning-based Control for Energy Management System

Rapid urbanization and disruptive technological progress in vehicular mobility have engendered a soaring growth in the automotive production in last two decades [1]. However, such growth may pose a detrimental threat to our planet by increasing emission of greenhouse gases (GHG)s. Forecasting the dreadful repercussions of uncontrolled GHG emissions on the planet's ecosystem, governments and automotive manufacturer decided to act upon the issue promptly through research collaborations. Electrification of automotive powertrain culminated as a promising field for combating against the aforementioned problem [2]. Based on the electrification factor, electrified vehicles can be broadly categorized into hybrid electric vehicle (HEV), plug-in-hybrid electric vehicle (PHEV), and electric vehicle (EV) [3], [4]. Partially electrified or HEVs are the most popular and also viable option because they are not encumbered with range anxiety and lack of charging infrastructure.

However, more energy sources incur additional complexity to the powertrain control. The complexity incurred due to electrification will be well-justified if all the energy sources are optimally controlled. A dedicated control system, known as energy management system (EMS), has become an integral part of the overall vehicle control module to carry out the aforementioned task. Fig.1.1 reiterates the dominance of Toyota Prius and other Toyota HEVs in U.S. market since the penetration of electrified powertrain into automotive market. Simple operation and effective control of Toyota hybrid system (THS) with two planetary gear (PG)-sets have been so successful that other original equipment manufacturer (OEM)s like Ford, Nissan licensed it

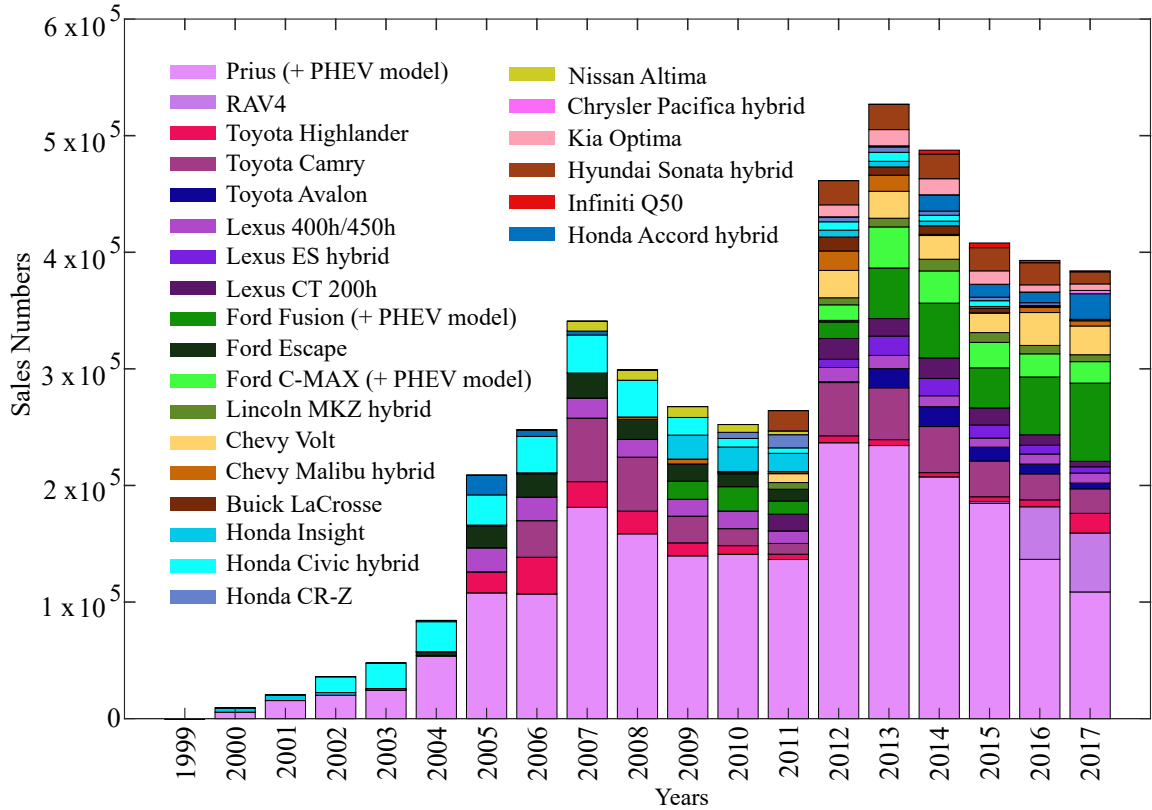


Figure 1.1: HEV sales trend in U.S. from 1999 to 2017 [5].

and used in their HEVs with a few modifications. In the meantime, the automotive industry has experienced an emergence of multiple electrified modes, introduced by General Motors (GM) in 2007, with two planetary gears and multiple clutches [6]. Alike automatic transmission, clutch-based mode-shift operation has been employed as the predominant field for encompassing so many electrified modes into a single hybrid electrified powertrain (HEPT) architecture. The two-mode electrified powertrain architecture coalesces the advantages of both parallel and power-split architectures in a single unit, making it even more versatile than power-split [7]. However, the mode-shift control becomes an added responsibility for the EMS whose control strategy needs to be even more competent than that of a series or parallel or even power-split



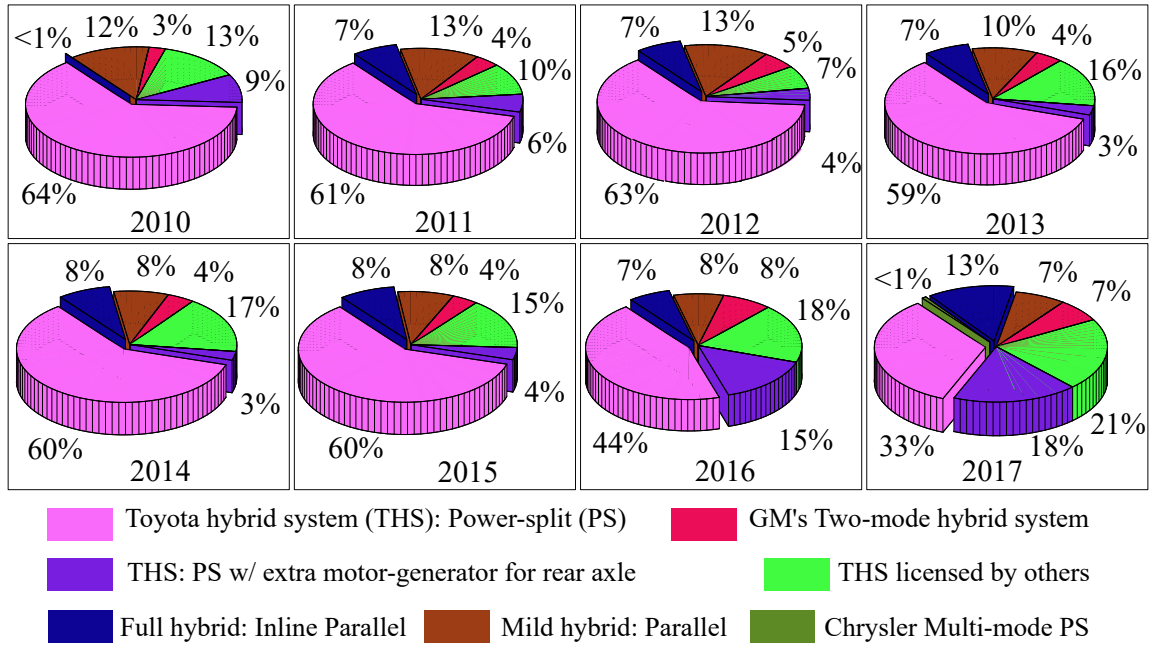


Figure 1.2: Sales trend of drivetrain architectures in U.S. from 2011 to 2017 [5].

architecture. Fig.1.2 and authors in [8] reveal that GM's two-mode hybrid system could not gain decent popularity even after complicated architecture topology with two power-split modes (i.e., input-split and compound-split), one parallel mode, two fixed gear modes, and prosperous attributes. On the other hand, the sales trend of Prius given in fig.1.1 attests the ingenuity of the simple architectural topology of third generation THS even without incorporation of any clutch operation. The stark contrast between the success of aforementioned two powertrains has ignited the research on advanced electrified powertrain architectures with multiple PG-sets and clutches.

The aforementioned reason has assisted in propagating researches in three distinct threads. First is, finding the optimal topology of powertrain architectures from a gargantuan design space comprising of architectures with two PG-sets, four (4) prime-movers (ICE, traction motor, generator, and vehicle output), and three (3) clutches allocated. Two PG-sets have six nodes where four prime-movers can be connected in

different permutation-combination. There are sixteen (16) feasible ways to connect six nodes of two PG-sets. Hence, we can have  $16_{C_3}$  ways of connecting three clutches in the two PG-set configuration. It is perceivable how huge would be the design space. Therefore, a systematic, swift, and effective screening process is highly recommended to find out optimal topology of powertrain architectures. Second, a systematic online simulation framework is indispensable to validate the performance of the optimal topology of architectures obtained from previous stage. The screening stage must be confined to offline simulation, executed by MATLAB<sup>®</sup>, to expedite the screening process. Since, online simulation framework for each powertrain architecture will need human intervention and cannot be automatized fully, screening of approximately a million design candidate will take forever if performance assessment of every design candidate is done through online simulation framework. Now, the online real-time simulation should incorporate dynamics of the powertrain components as well as the clutch dynamics during mode-shift events. Third and the last thread can be seen as the most important among all these because of its association with articulation of EMS. According to authors in [4] and [9], an intelligent EMS should possess a few indispensable qualities in order to qualify for online simulation framework for an electrified powertrain with multiple modes and these are enumerated as follows:

1. The EMS should be real-time implementable with less computational burden.
2. It should be independent of human intervention once the vehicle is deployed on road.
3. It should yield near-global optimal performance in terms of fuel consumption in real-world driving missions.

4. It should be continuously or periodically updated during real-world driving situations.
5. Last but not the least, considering it's association with a powertrain with multiple electrified modes, it should be equipped with coordinated control during mode-shift events.

Now, a brief discussion will be presented here on the available control strategies which can be explored for an EMS having aforementioned qualities. The global optimal control strategy obtained through Dynamic Programming (DP) is excluded from the following list considering the fact that DP cannot be implemented in the real-time processors of vehicle control unit (VCU).

1) Rule-based or premeditated control strategy. The premeditated control strategy is comprised of some calibrated look-up tables or driving situation-specific set of rules which can yield global optimal only for apriori drive cycles. Being extracted from global optimal control for a set of apriori drive cycles, the premeditated control strategy is stupendous in generating global optimal control in real-time real-world driving. However, the real-world driving has to be a member of the aforementioned set of apriori drive cycles. Upon encountering an unknown drive cycle, premeditated control strategy cannot even guarantee near-optimal performance. Once deployed on road, the premeditated strategy that is embedded on the VCU, cannot be altered through human intervention during vehicle's running life.

2) Periodically-updating stochastic dynamic programming (SDP) based control strategy. SDP is an useful way of tackling uncertainty in power demand during real-world driving scenario [10]. While deterministic DP is applied on a single drive cycle, SDP is applied to an ensemble of drive cycles [11]. SDP yields a multi-dimensional look-up

table comprising of the control strategy which is applicable for the entire ensemble of drive cycles. However, the SDP is computationally demanding and takes even more time than deterministic DP and hence cannot be implemented in VCU for real-time application. Still, the control strategy engendered by the SDP can be used during real-world driving situations due to fact that the control strategy is not only tailored just for one drive cycles, but also for a many more variation of driving situations. Now, while deployed on road, the VCU can collect driving data and the control strategy, yielded from SDP, can be periodically updated based on the recent trend of driving data. It is noteworthy to mention that the periodic computation of SDP can be either executed vehicle onboard if VCU is equipped with powerful processor or executed in the cloud computing facilities.

3) Real-time implementable Model predictive control (MPC). The prevalent principle of MPC is to predict disturbances to the system over a receding future time-window and manipulate the control variables to optimize the system performance. For EMS of HEVs, different methods have been witnessed in literature while implementing MPC as a real-time control strategy. Power-demand of the vehicle is the predominantly used disturbance variable since it can encompass effect of many disturbance variables such as vehicle speed, vehicle acceleration, road grade, wind speed, etc. Different tools such as “Markov Chain Model” [12] and exponential predictor [13, 14] are of the available for predicting power-demand. Optimal control methods such as DP [13, 14] and Quadratic Programming (QP) [12] can be employed to optimize system performance within the finite time-window of future. Despite their enticing application as a real-time updating near-optimal control strategy for the EMS of an HEV, their success relies highly on the accuracy of prediction.

4) Equivalent Consumption Minimization Strategy (ECMS) with a real-time adaptive rule for the equivalence factor is strong contender as the desired control strategy [15]. ECMS has been widely accredited as a cogent solution for the EMS of an ePT despite ECMS's strong reliance on equivalence factor. ECMS has been proven as the reincarnation of the Pontryagin's minimum principle (PMP) [16] in optimal control of energy management problem. If driving mission is known in advance, ECMS can yield optimal control having close proximity to DP [17]. Although ECMS cannot achieve global optimality in the absence of prior knowledge of the drive cycle during real-world driving conditions, it can achieve near-global optimality along with charge sustenance if equivalence factor can be varied judiciously [18]. Authors in [19] and [20] concurrently introduced the concept of adaptive ECMS (A-ECMS) where the equivalence factor was updated periodically based on past trend of vehicle speed and prediction of future driving mission. Later, adaptation of equivalence factor has been executed through different methods such as tracking of state-of-charge (SOC) reference value [21, 22], prediction of future power demand [23], and recognition of current driving pattern [24].

Looking at the strength and drawbacks of the aforementioned available control strategies, author of this report has decided to delve into the fundamentals of sequential decision making to search for an apposite control strategy for the EMS of a multi-mode ePT. On the other hand, reinforcement learning (RL) has been emerging as a lucrative control strategy for optimizing controllable physical systems in real-time since the intrusion of artificial intelligence into the sector of optimal control. RL mathematically models an intertwined psychological and neurological behavior, exhibited by humans and a few animals, for continuously optimizing decision making





Case	State variables (number of discretization)			Action variables (number of discretization)			Approx. number of cells required in the Q-table		
2 state 1 control	$\nu_{veh}(m/s)$ (20)	$SOC(\%)$ (20)		$\tau_{ice}(N-m)$ (20)			$20 \times 20 \times 20 = 8000$		
3 state 2 control	$\nu_{veh}(m/s)$ (20)	$SOC(\%)$ (20)	$Road\ grade$ (5)	$\tau_{ice}(N-m)$ (20)	$\omega_{ice}(rpm)$ (20)		$20 \times 20 \times 5 \times 20 \times 20 = 400000$		
3 state 2 control	$\nu_{veh}(m/s)$ (40)	$SOC(\%)$ (20)	$Road\ grade$ (10)	$\tau_{ice}(N-m)$ (40)	$\omega_{ice}(rpm)$ (40)		$40 \times 20 \times 10 \times 40 \times 40 = 12800000$		
4 state 3 control	$\nu_{veh}(m/s)$ (40)	$SOC(\%)$ (20)	$Road\ grade$ (10)	$Pow_{dem}$ (40)	$\tau_{ice}(N-m)$ (40)	$\omega_{ice}(rpm)$ (40)	$Gear$ (5)	$40 \times 20 \times 10 \times 40 \times 40 \times 40 \times 5 = 2560000000$	

Figure 1.3: Curse of dimensionality .

capability based on past experiences to substantiate gain in future [25], [26]. Authors in [27] presented the first use of RL in articulating the control strategy for EMS with a ADVISOR-based vehicle simulation model.

## 1.2 Motivation

Although RL’s emergence as a control strategy for the EMS seems to solve most of the drawbacks mentioned above, traditional RLs based on discrete-time domain faces a few legitimate hindrances at the implementation level. RL-based controllers are prone to the “curse of dimensionality”. The “curse of dimensionality” might not appear with fewer variables or a bigger grid size for discretization. Nevertheless, the computational burden increases exponentially with either an increased number of variables or decreased grid size. Fig.1.3 has shown the approximate number of cells required in a Q-table with traditional discrete RL methods such as Q-learning and SARSA.

Although traditional RL algorithms are model-free, they can learn autonomously, and they achieve convergence, they take much time to converge to a near-optimal performance. Fig.1.4 depicts that both Q-learning and SARSA steadily converge but

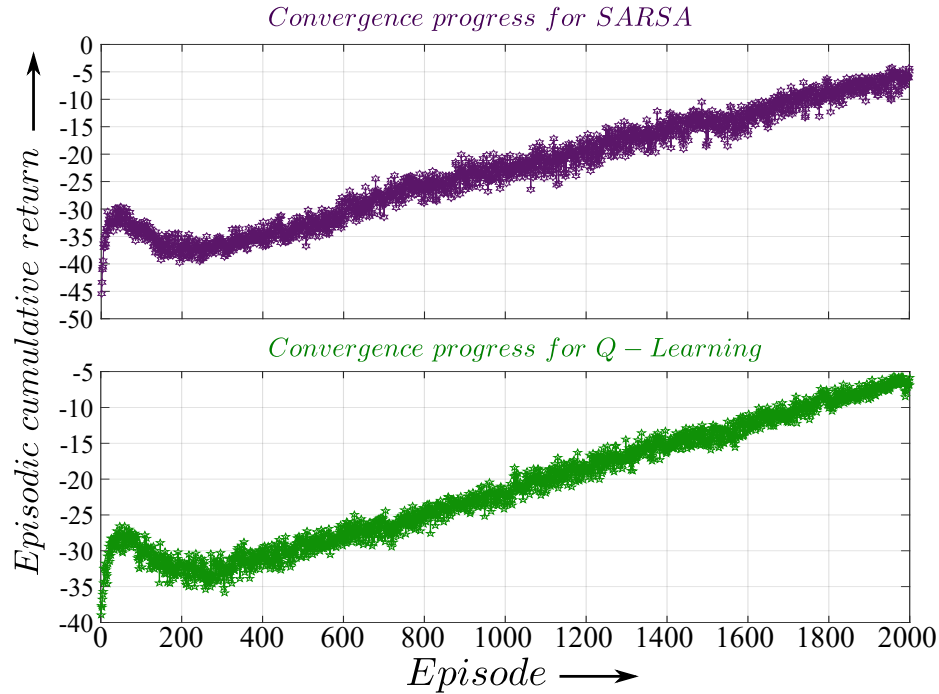


Figure 1.4: Sluggishness of convergence for traditional RL algorithms with discrete-time domain.

with much sluggishness. Moreover, in a MATLAB-based implementation, it took approximately two days of training to reach this level of convergence.

However, in real-world driving, the EMS inevitably faces legitimate challenges. The real-world driving scenario is comprised of several parameters. 1) Knowledge of the destination, 2) Vehicle’s speed, 3) Vehicle’s acceleration, 4) Slope of the road, 5) Ambient temperature, 6) Wind speed, 7) Traffic situation, 8) State of charge of the battery, 8) Distance from the following charging points and so on. Amongst these, many parameters are unknown in real-time during real-world driving. For example, even though the destination is selected, there could be multiple routes to reach the destination, as shown in Fig.1.5.a). Another example is the inaccurate estimation of traffic ahead, as shown in Fig.1.5.b). Being such a versatile control strategy for

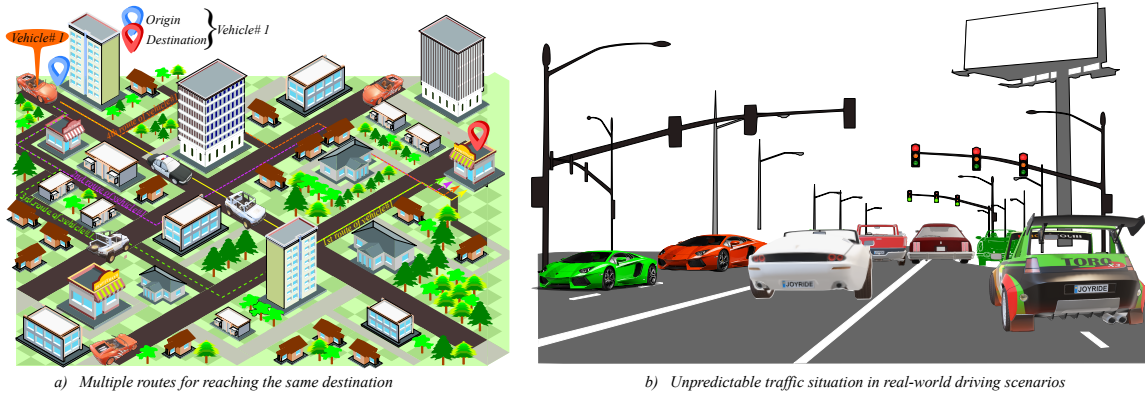


Figure 1.5: Unpredictable parameters such as traffic situations and driving routes to the destination during real-world driving scenario.

solving the energy management problem of an electrified powertrain for unknown drive cycles, RL should not be discarded only because of the aforementioned challenges. This becomes the motivation for eliminating those hindrances and establishing the RL’s superiority as a prospectus on-board EMS control strategy.

### 1.3 Objectives and Contributions

The main objective of this thesis is to articulate a real-time updating and near-optimal control strategy for the EMS of a multi-mode ePT. The articulated framework has the potential of being deployed in a real multi-mode HEV for yielding near-global optimal control in unfamiliar driving situations. While articulating such an ingenious EMS it is perceived that mode-scheduling control for a multi-mode ePT necessitates special attention compared to normal ePT. Hence, this thesis can be considered as a constellation of three main contributions as follows:

1. A benchmark instantaneous optimization-based strategy, named ECMS, is implemented for a multi-mode ePT. This EMS control strategy is articulated for



comparing newly developed strategies in this thesis. In a pursuit for new real-time near-optimal control, this thesis proposes a loss-minimization strategy for powertrain’s loss minimization.

2. An asynchronous actor-critic agent is employed as a real-time implementable near-optimal EMS strategy, which can be deployed for real-world driving scenario with online updating feature.
3. A low-frequency coordinated control is articulated for correct mode-scheduling of multi-mode ePT to corroborate its necessity for an inertia-based dynamic powertrain model.

## 1.4 Summary of Thesis

The chapter 2 starts with journey of this thesis by highlighting the key role of powertrain electrification in ameliorating the adverse effects of greenhouse gas emissions. Without wasting much time, chapter 2 delves into the fundamentals and classification of electrified powertrain, and creates a smooth avenue for the reader into the domain of powertrain electrification. Working principle of different powertrain configurations becomes elucidated in chapter 2 and it serves as the perfect foundation for acquiring advanced knowledge in powertrain electrification. chapter 2 also incorporates a brief description of some existing commercial trans-axles such as THS, “*GM’s two mode hybrid system*”, and Chrysler’s Si-EVT.

Powertrain electrification augments one or more than one degree-of-freedom in the

entire propulsion system to enable an efficient control. However, the extra degree-of-freedom comes at the cost of added complexity into the control system. Powertrain electrification brings more cardinal components into a cohesive environment, whose benefits predominantly depends on the optimal control of the entire powertrain system. An exclusive control system known as energy management system is indispensable for extracting all the benefits of powertrain electrification. Articulation of energy management system can be as simple and nimble as rule-based or as complex and computationally expensive as dynamic programming. A detailed review of the existing EMS control strategies with the future trend of strategies is presented in chapter 3. chapter 3 also presents a tentative evolution and popularity trend of the control strategies, and indicates towards possible research gap in the EMS strategy. Since the main stem of this thesis is energy management system and the thesis develops around it, chapter 3 carries a special importance in the entire thesis.

Since the lucrative aspect of the thesis is the implementation of a real-time optimal energy management system, it is recommended to revisit an existing optimization-based real-time control from literature. ECMS is a well-accredited local optimal control, which has been proven to be performing near-global optimally if drive cycle is known in advance. Hence, ECMS is considered to be the benchmark control in this thesis and the thesis will progress towards finding a novel control strategy, which can give near-optimal performance in real-time under zero knowledge of the drive cycle ahead. Most importantly, chapter 4 elucidates all implementation phases of the benchmark ECMS for EMS. This benchmark strategy will be compared to all newly proposed strategies in this thesis.

chapter 5 proposes the first of the newly professed control strategies in the form of

loss-minimization of the entire powertrain at every time-step. The control strategy is similar to ECMS, and is proposed to minimize the overall powertrain loss unlike ECMS which focuses on just maximizing ICE's efficiency. The loss minimization strategy (LMS) is articulated with an ambition of outperforming optimal ECMS, given that LMS does know the drive cycle apriori.

chapter 6 serves as the prelude to intricate and advanced reinforcement learning-based control strategies proposed in this thesis. Before delving into the advanced application of RL algorithm in optimizing control policy in real-time, fundamentals of RL algorithm is delineated in this chapter. Markov decision problem is a widely-accredited mathematical model of sequential decision and it serves as the foundation of RL algorithms. Hence, this chapter first delineates the components and characteristics of Markov decision problem, and then introduces RL algorithm as an adroit tool to solve Markov decision problem. The chapter culminates with the detailed implementation of Q-learning as an energy management strategy for Toyota Prius.

After chapter 6 the stage is all set to profess the main contribution of this thesis, i.e., application of asynchronous actor-critic agent as the EMS strategy for real-world driving in chapter 7. The framework emulates the real-time interaction between the agent and the electrified powertrain and shows how efficient the agent can yield near-global optimal control for unfamiliar real-world driving scenarios. This chapter elucidates the asynchronous actor-critic agent's effectiveness as a prospective real-time and real-world controller for a multi-mode electrified powertrain's energy management system. This chapter posits the key novel contribution of this article.

While discussing different EMS implementation for a multi-mode ePT in online simulation platform it has been understood that a steady-state powertrain model can be

controlled with a steady-state EMS controller. However, a steady-state powertrain model will not capture all the transient dynamics of an actual physical powertrain. Hence, an inertia-based dynamic powertrain model is indeed necessary for better emulation of an actual powertrain. Consequently, a dynamic coordinated control is indispensable for correct mode-scheduling control. Thus, as a second contribution of this thesis, a low-frequency dynamic coordinated control is proposed in chapter 8.

Chapter 9 concludes the thesis and discusses the future work for further improvements in EMS and dyno validation of the newly proposed strategy's performance. In a nutshell, thesis shows a tip of an iceberg as far as the EMSs for electrified powertrains are concerned. The thesis ends with avenues in multiple directions of research gap, i.e., development of high-frequency EMSs for real-time application with high-frequency coordinated control, effect of high-fidelity powertrain model on the EMS development, online training of a reinforcement learning agent working as a EMS controller, practical consideration while implementing an online trainable reinforcement learning agent on a real-time embedded controller, etc.

## **Chapter 2**

### **Electrified Powertrain:**

### **Fundamentals and Classification**

## 2.1 Introduction

In this chapter, the fundamentals of ePT will be presented along with the classification of ePTs based on the degree of electrification. This chapter will delineate the rudimentary architecture configuration along with state-of-the-art multi-mode architectures and their modeling techniques.

## 2.2 Requirement of Electrified Powertrains

The never-ending quest for a safer, more comfortable, smarter, and more economical mode of transportation has started since the official birth of modern car (Benz Patent-Motorwagen) in the year of 1886. Sales and production figures of automotive industry became one of the key factor in determining economy of a country and soon the automotive industry became a major contributor to the global economy [28]. Number of registered new vehicles across the world can be perceived as an apposite metric of global economy and Fig.2.1 corroborates the proportional behavior between global economy and automotive sales numbers [1]. However, such growth may pose a detrimental threat to our planet by increasing emission of GHG. As per U.S environmental protection agency (EPA), transportation sector contributes to the highest share (29.1%) of national GHG emissions [29]. Fig.2.2 depicts the growth in GHG emissions only caused by transportation sector in U.S. According to [30], transportation is the second largest sector contributing to GHG emissions in Canada. Fig. 2.3 shows that soaring trend of GHG emissions in transportation sector is caused by mainly due to passenger light trucks and freight trucks which mostly have conventional fuel-powered powertrains. Forecasting the repercussions of uncontrolled GHG

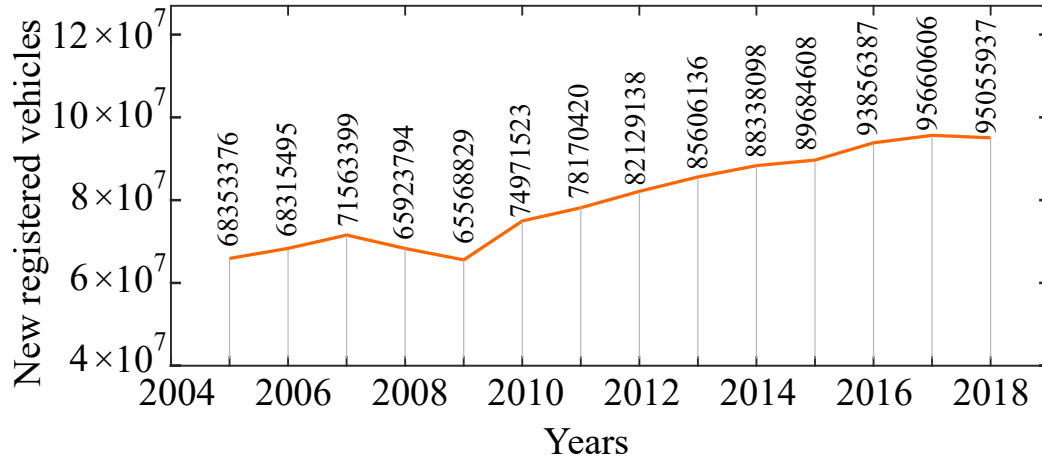


Figure 2.1: Automotive sales statistics in last one and a half decades [1].

emissions on the planet’s ecosystem, governments of most of the developed nations enhanced the stringency of their environmental regulations in 1990s. United nations framework convention on climate change (UNFCCC) had set emission reduction targets in 1997 for all the developed nations, which contribute to lion’s share of the global GHG emissions, through an unanimous agreement known as the Kyoto Protocol. Scholars have proffered their argument that such stringency is actually beneficial for the OEMs by prompting them to invest more in research which will lead the OEMs to innovations [31]. Two case studies done in [32] corroborate the above-mentioned argument by showing how Japan’s stringent environmental regulations had paved the way for automotive OEMs like Toyota, Nissan, and Panasonic manufacturing innovative and eco-friendly electrified vehicles like Toyota Prius, Nissan Leaf, and Panasonic Eco Navi.

Bestowed with rapid technological advancements in the automotive electrification, the future of automobile industry will experience an inevitable paradigm shift from liquid fuel based powertrain to ePT. HEV is the first step towards that energy-efficient

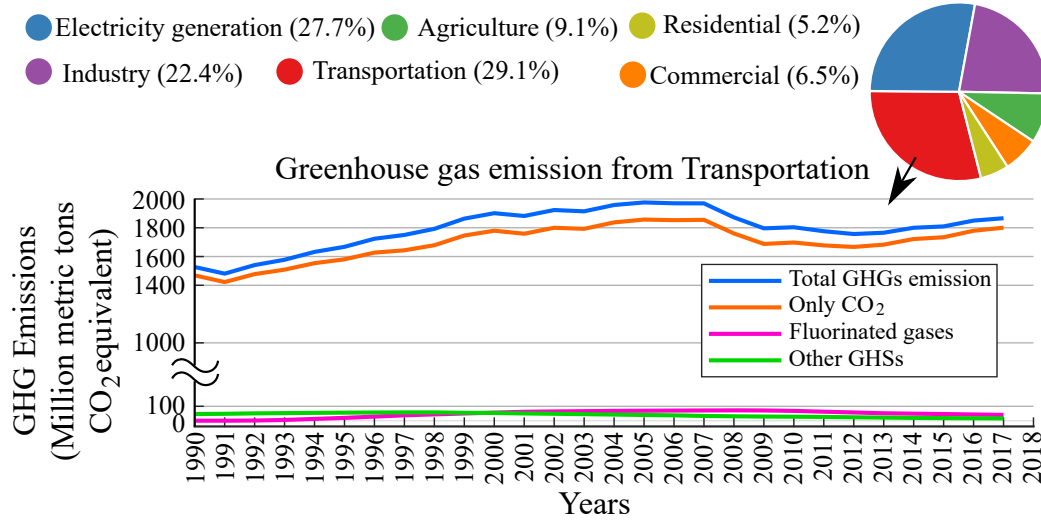


Figure 2.2: U.S. GHG emissions by economic sector, 2017 and GHG statistics for transportation sector [29].

era [2]. Introduction of electrification brings an extra dimension in the powertrain control with the provision of more than one energy source [33]. The primary objectives of transition from ICE based powertrain to ePT is reduction of both fuel consumption and tailpipe emission whereas, a few of the secondary objectives are improvement of drivability, handling, and acceleration performance.

## 2.3 Classification Electrified Powertrain Architectures

The articulation of EMS relies strongly on architectures of ePT. Hence, an overview of the available ePT architectures is indispensable before proceeding to the next sections detailing about different classes of EMSs. ePTs can be broadly clustered into micro hybrid, mild hybrid, strong hybrid, plug-in hybrid, and fully-electric powertrains [34, 35]. Based on the electrification factor and location of the cardinal components,



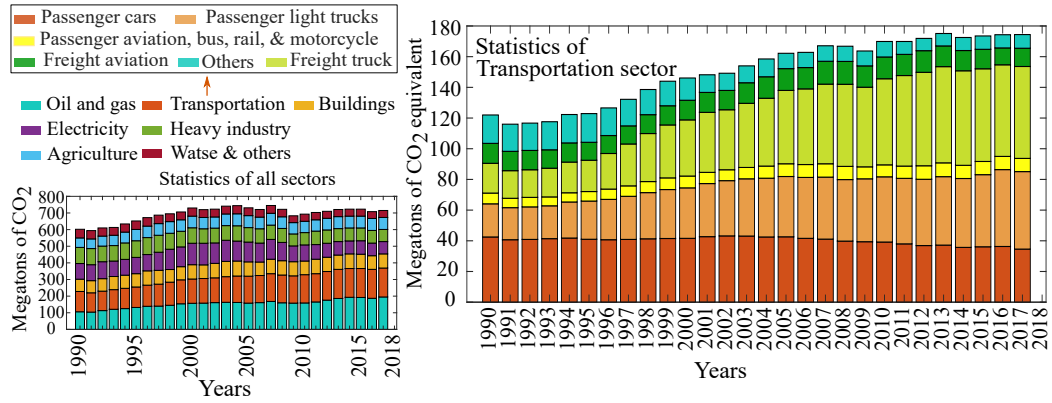


Figure 2.3: Variation of Canada’s GHG emissions since 1990 to 2018 in different economic sectors and GHG statistics for transportation sector [30].

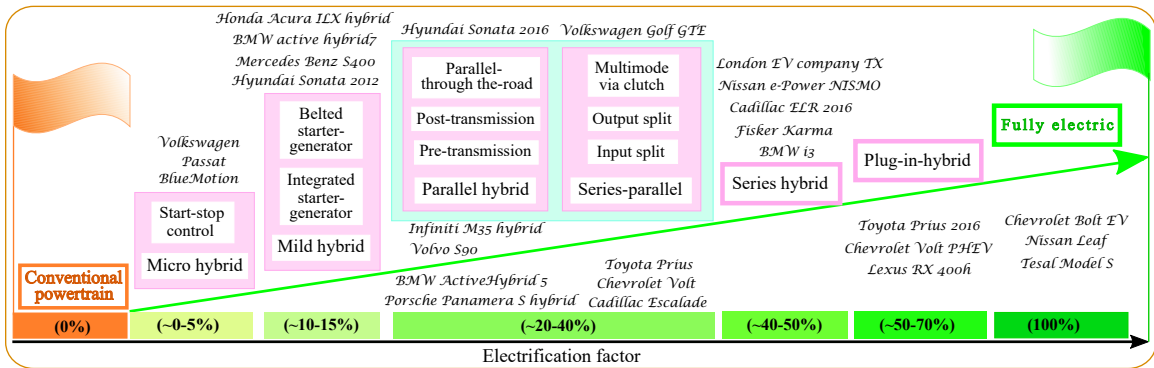


Figure 2.4: Hierarchical classification of architectures of electrified powertrains.

hybrid ePTs can be categorized into series, parallel, and series-parallel architectures [36], [37]. In order to combine the advantages of series and parallel architectures and to discard disadvantages of both, new class of architecture named as series-parallel, has been conceptualized [38]. The concept of series-parallel can be embodied either with the assistance of clutch operation or with a power-split device (PSD) [39]. A general hierarchy of architectures of ePTs is shown in Fig. 2.4. This article focuses only on the architectures of ePTs which include both ICE and electric motor/generator (EMG)s.

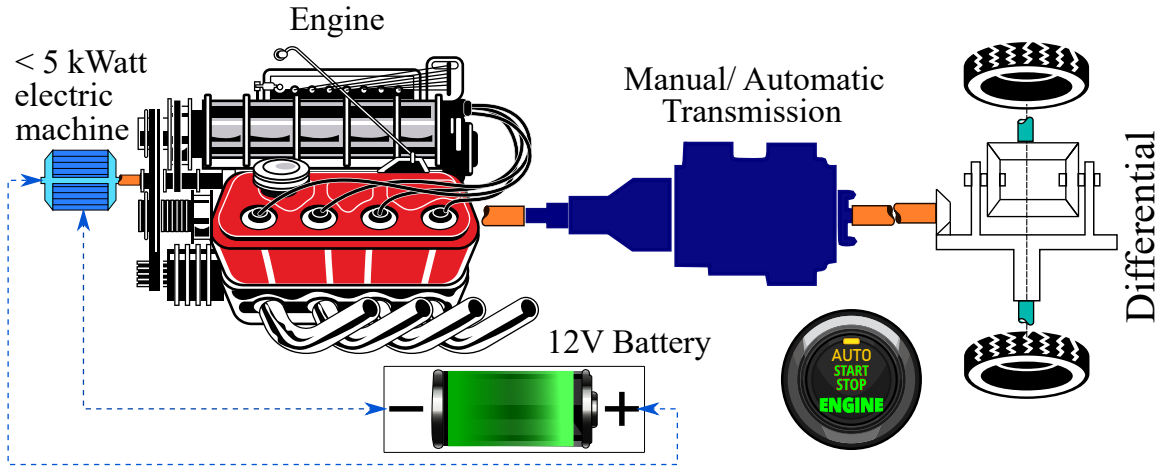


Figure 2.5: Architecture of a micro hybrid electric powertrain with less than 5% electrification factor.

### 2.3.1 Micro hybrid Architecture

Micro hybrid architecture, as depicted in Fig.2.5 shares almost the same powertrain configuration as traditional liquid-fuel based powertrain. Incorporation of a low-powered EMG in the form of starter generator is common in conventional powertrain. However, the VCU of a micro hybrid vehicle is equipped with a dedicated control system which automatically cut offs the fuel supply to the ICE as soon as vehicle comes to rest upon application of brake. The control system is known as "auto start-stop" strategy. The strategy again automatically cranks the ICE through the small EMG as soon as driver takes her/his foot off the brake pedal. This "auto start-stop" strategy helps in curtailing fuel consumption a few percentage as compared to conventional powertrain. Still, a 12V battery is used in micro hybrid architecture.

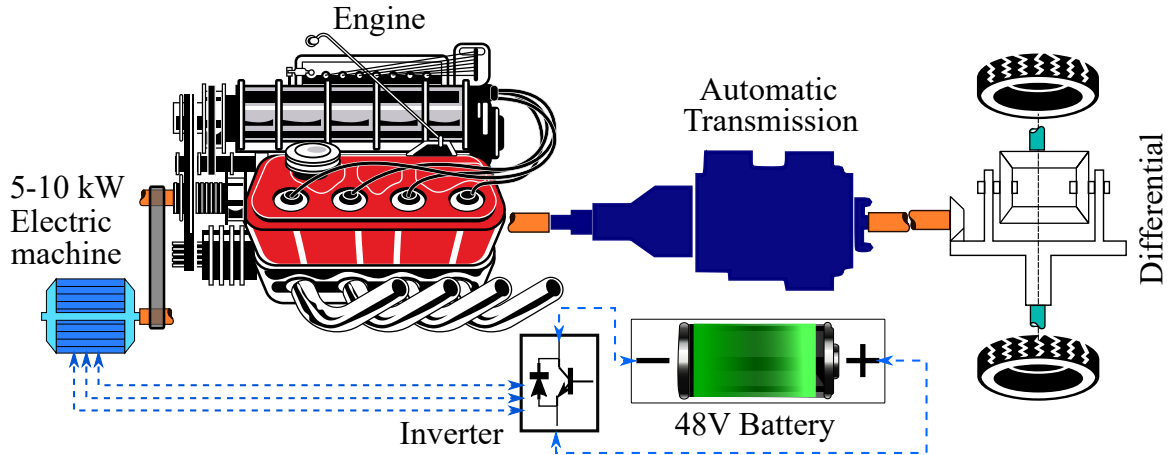


Figure 2.6: Architecture of a mild hybrid electric powertrain with (10%-15%) electrification factor.

### 2.3.2 Mild hybrid Architecture

Mild hybrid architecture retains the same architecture of micro hybrid but includes higher battery capacity and higher EMG power unlike micro hybrid. There is no fixed regulation for categorizing mild hybrid architecture. Either a 42V or a 48V battery used instead of 12V battery in mild hybrid architecture. There two types of mild hybrid architectures available based on the position of the 5-10kWatt EMG. The first type is shown in Fig.2.6 where the EMG is placed before the ICE either as an "integrated starter-generator" or "belt-driven starter-alternator". The tasks of the EMG in a mild hybrid powertrain are to crank the ICE and to assist the ICE during high power-demand. Since the EMG is not connected to the wheel through the transmission, recuperation of energy during braking cannot be availed in this category of mild hybrid architecture. Vehicles like Honda Acura ILX hybrid, Mercedes Benz S400, and Hyundai Sonata 2012 had this category of mild hybrid architecture.

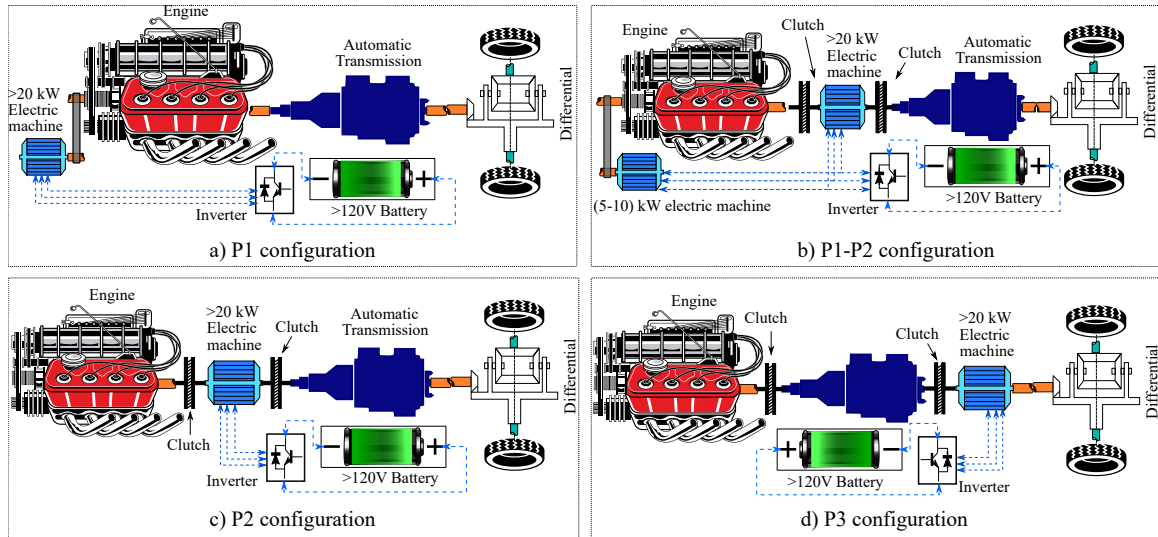


Figure 2.7: Different configurations of a P1 hybrid electric powertrain architecture with (20%-40%) electrification factor.

### 2.3.3 Parallel hybrid Architecture

Parallel hybrid architecture brings decent amount of electrification in the powertrain. The EMG is capable of delivering at least 20kWatt power and therefore, the battery also has to be bigger in capacity. The battery voltage is no lesser than 120V. The key characteristic of parallel architecture is to combine the powers of both ICE and EMG/s and delivers to the wheel. Another key aspect of parallel architecture is the mechanical coupling among ICE, EMG, and the wheel. Now, there are a few ways to place the EMG and combine the powers of ICE and EMG and then propagate to the wheels. Based on these ways, parallel architecture can be further sub-categorized into five sub-classes, as shown in Fig.2.7 and in Fig.2.8. The P4 parallel configuration is also known as "parallel through road" having the EMGs directly attached to the wheel. The ICE is not mechanically decoupled from the wheel in any of the parallel configurations. Hence, the ICE speed is always governed by the vehicle's speed. This

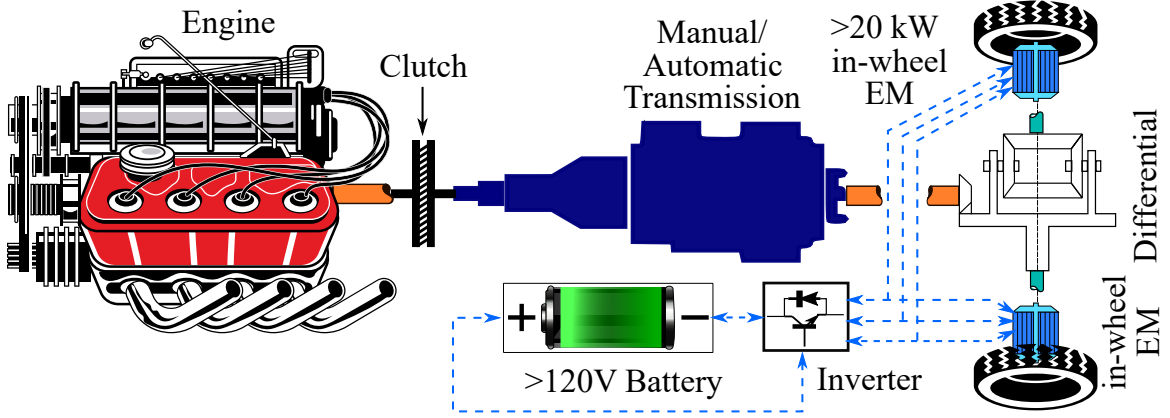


Figure 2.8: Architecture of a P4 hybrid electric powertrain with (20%-40%) electrification factor.

is one of the drawbacks of parallel architectures. However, the engine torque can be controlled independently with respect to the vehicle's power demand due to the presence of the EMG/s. Typically, an ICE in a conventional powertrain doesn't have any degree-of-freedom (DOF). An ICE leverages one DOF in a parallel architecture. The parallel architecture can only enable torque-split between ICE and the EMG/s. The generic speed and torque relationships among transmission output, ICE and EMG is resented in Eq.2.3.1.

$$\tau_{ice} + \tau_{EMG} = \tau_{out} = \left( \frac{\tau_{wheel}}{final\ drive\ ratio} \right) \frac{1}{N^{th}\ gear\ ratio} \quad (2.3.1a)$$

$$\omega_{ice} = \omega_{wheel} \times final\ drive\ ratio \times N^{th}\ gear\ ratio \quad (2.3.1b)$$

$$\tau_{ice} = \frac{\tau_{out}}{\xi}; \text{ where } \xi \text{ is the torque – split ratio} \quad (2.3.1c)$$

$$Pow_{batt} = \frac{Pow_{EMG}}{\eta_{EMG}} [for\ Pow_{EMG} > 0] \text{ or } Pow_{EMG} \times \eta_{EMG} [for\ Pow_{EMG} < 0] \quad (2.3.1d)$$

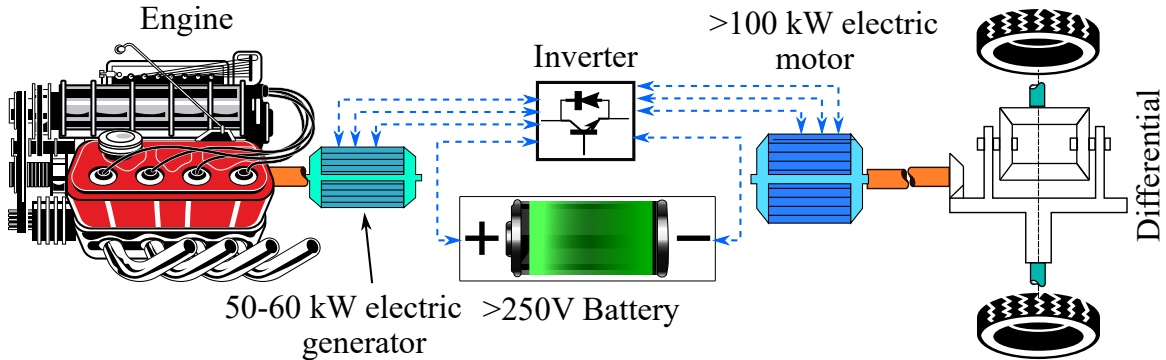


Figure 2.9: Architecture of a series hybrid electric powertrain with (40%-50%) electrification factor.

### 2.3.4 Series Hybrid Architecture

The powertrain configuration of series hybrid architecture is shown in Fig.2.9. The main attribute of series hybrid architecture is the mechanical decoupling of ICE from the wheel. An EMG, working as a electric generator, is directly connected to the ICE and it recharges the battery. The generator converts the mechanical energy churned out of the ICE. The vehicle is solely propelled with a high-powered EMG working as a traction motor. The traction motor draws energy from battery during traction and replenishes the battery during regenerative braking. Since the ICE is mechanically decoupled from the wheel, its speed can controlled independent from the wheel speed. The series hybrid architecture is also known as "range-extended electric vehicle". The ICE coupled with the generator is also known as the auxiliary power unit (APU) of the vehicle. The APU kicks-in when SOC of the battery depletes beyond its permissible limit. The series configuration enables two-DOF in terms of powertrain control. Both ICE speed and ICE torque can be controlled with respect to any vehicle's power demand. The generic power relationship among the key components of series hybrid

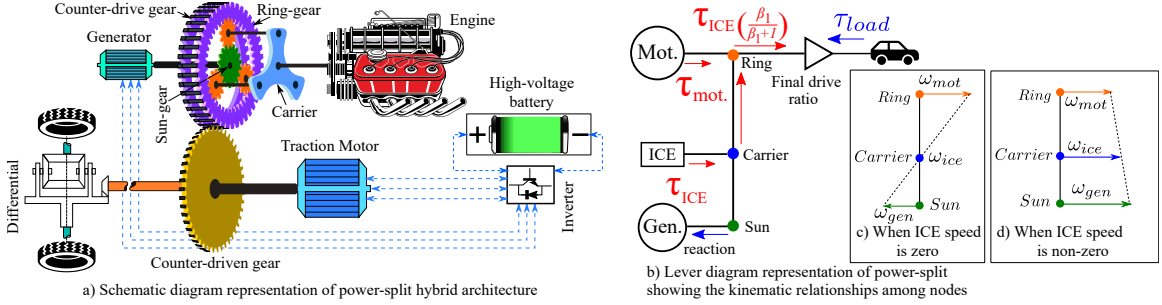


Figure 2.10: Architecture of a power-split hybrid electric powertrain with (20%-40%) electrification factor.

architecture is given in Eq.2.3.2.

$$\eta_{mot} * Pow_{mot} = \tau_{out} = \left( \frac{\tau_{wheel}}{final\ drive\ ratio} \right) \frac{1}{N^{th}\ gear\ ratio} \quad (2.3.2a)$$

$$Pow_{mot} = \eta_{mot} (Pow_{gen} + Pow_{batt}) \quad (2.3.2b)$$

Despite the provision of two-DOF control, series hybrid is poised with a conspicuous drawback of too many number of power conversion and energy loss during these conversions. The mechanical energy is converted into electrical energy in the APU. Subsequently, electrical energy is stored into battery and again that stored electrical energy is converted into mechanical energy through the traction motor.

### 2.3.5 Power-split Hybrid Architecture

To combine the advantages of both series and parallel configurations and to address most of the drawbacks of these two configurations, a new configuration was conceptualized and deployed as a power-split configuration. The power-split concept transpired into a tangible architecture only with an assistance of a single PG-set. A PG-set is a versatile and ingenious mechanical device used for coupling torque and

speed of the prime-movers connected to the three nodes of a PG-set.

A typical power-split configuration is shown in Fig.2.10, where the ICE, generator, and traction motor are connected to the planet carrier, sun gear, and ring gear of the PG-set. The vehicle load coming from the wheel is also connected to the ring gear the PG-set. One of the lucrative characteristics of PG-set is its capability of enabling infinite number of gear-ratio between input (carrier) and output (ring gear). Unlike parallel configuration, ICE speed can be independently controlled for any given positive vehicle's speed. Unlike series configuration, ICE can deliver directly power to the wheel. Most importantly, there is no need of any transmission which would have contributed loss of energy. The capability of providing infinite number of gear-ratio between input-output with the assistance of EMGs gives the power-split configuration another identity named electrically continuous variable transmission (e-CVT). The kinematics and dynamics associated with the power-split configuration, shown in the Fig.2.10, is given in Eq.2.3.3.

$$\tau_{sun} = \tau_{gen} - J_{gen}\ddot{\theta}_{gen} \quad (2.3.3a)$$

$$\tau_{carrier} = \tau_{ice} - J_{ice}\ddot{\theta}_{ice} \quad (2.3.3b)$$

$$\tau_{ring} = \frac{N_{ring}}{N_{sun}}(\tau_{sun}) ; \tau_{ring} = -\frac{N_{ring}}{N_{ring} + N_{sun}}(\tau_{carrier}) \quad (2.3.3c)$$

$$\tau_{out} = \tau_{ring} + J_{out}\ddot{\theta}_{out} + \tau_{mot} - J_{mot}\ddot{\theta}_{mot} \quad (2.3.3d)$$

$$\tau_{out} = \left( \frac{\tau_{wheel}}{final\ drive\ ratio} \right) \frac{1}{N^{th}\ gear\ ratio} \quad (2.3.3e)$$

$$\omega_{gen} = \omega_{ice} \left( \frac{N_{ring} + N_{sun}}{N_{sun}} \right) - \left( \frac{N_{ring}}{N_{sun}} \right) \omega_{out} \quad (2.3.3f)$$

$$\omega_{mot} = \omega_{out} \quad (2.3.3g)$$





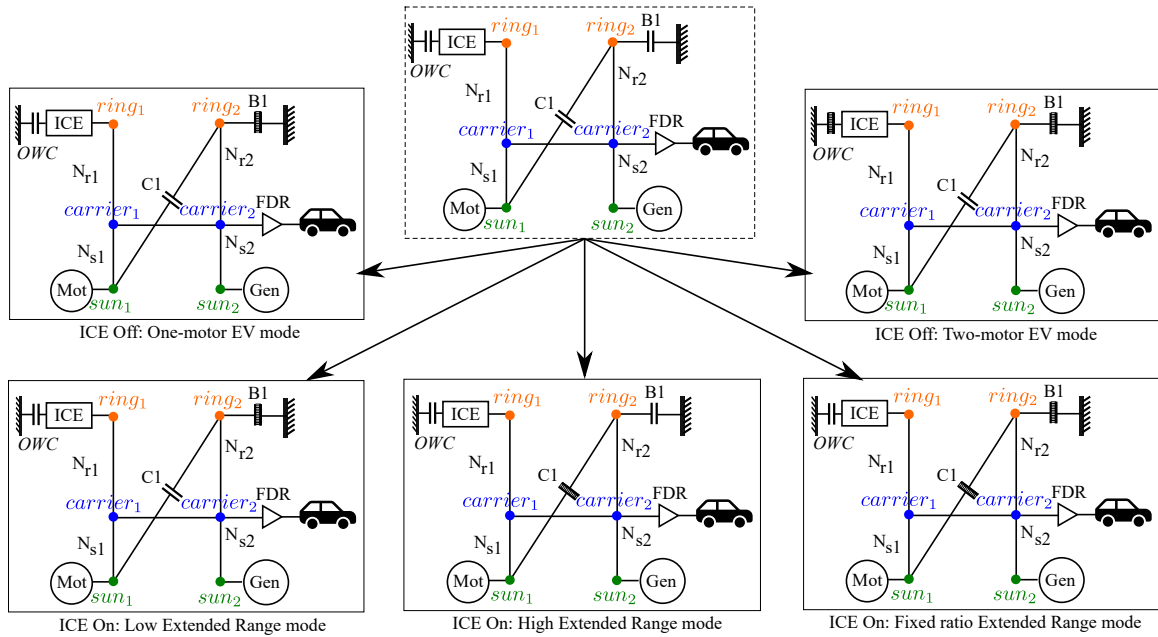


Figure 2.12: Schematic diagram of advanced “Two-mode hybrid system” with two PG-sets.

based GM’s two-mode hybrid system” is furnished in Tab.2.1. This multi-mode architecture topology facilitates four electrified modes. The summary of clutch engagement and disengagement status required for different modes in “Double-PG-set based GM’s two-mode hybrid system” is furnished in Tab.2.2. This topology facilitates five electrified modes. A few of the application vehicle equipped with the aforementioned multi-mode ePTs are Chevrolet Volt (first generation and second generation) and 2016 Chevrolet Malibu.

### Chrysler’s Single-Input Electrically Variable Transmission

Recently Fiat-Chrysler has introduced a modified version of THS with three electrified modes [42]. The detailed schematic diagram along with the lever diagram representation of all the three modes is shown in Fig.2.13. Three distinct electrified

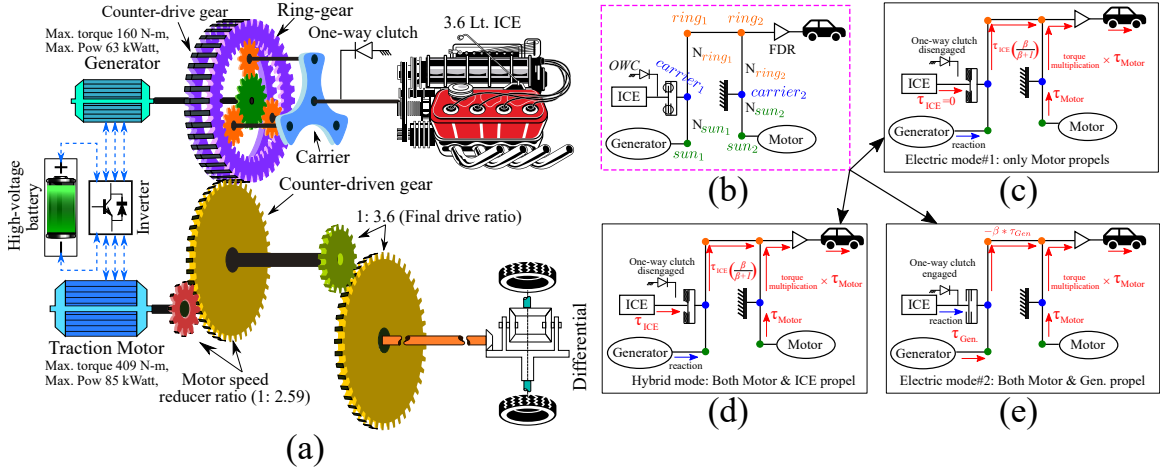


Figure 2.13: Schematic diagram of Single-Input Electrically Variable Transmission with two PG-sets and the modes facilitated by it.

modes, i.e., single-motor EV mode, dual-motor EV mode, and a hybrid-electric mode are facilitated by the aforementioned configuration. As depicted in the Fig.2.13, ICE is attached to the planet carrier of the PG set through an one-way clutch (OWC) (sprag clutch) which allows the ICE to rotate only in the one direction. The generator is attached directly to the sun gear, and the traction motor is attached to the ring gear through a speed reduction gear and a counter-driven gear. The single-motor EV mode and the hybrid-electric mode share the same powertrain dynamics as shown in Fig.2.13.(c) and Fig.2.13.(d) respectively and it is expressed through a

set of kinematic relations as given by Eq.2.3.4.

$$J_{out_{eq}}\ddot{\theta}_{out} - \tau_{out} = \frac{\beta_1}{(\beta_1 + 1)} * \{\tau_{ice} - J_{ice_{eq}}\ddot{\theta}_{ice}\} \quad (2.3.4a)$$

$$+ \beta_2 * \{J_{mot_{eq}}\ddot{\theta}_{mot} - \tau_{mot}\}$$

$$\tau_{ice} - J_{ice_{eq}}\ddot{\theta}_{ice} = (\beta_1 + 1) * \{J_{gen_{eq}}\ddot{\theta}_{gen} - \tau_{gen}\} \quad (2.3.4b)$$

$$(\beta_1 + 1) * \ddot{\theta}_{ice} = \beta_1 * \ddot{\theta}_{out} + \ddot{\theta}_{gen} \quad (2.3.4c)$$

$$\ddot{\theta}_{mot} = -\beta_2 * \ddot{\theta}_{out} \quad (2.3.4d)$$

The ICE is not cranked at all in the two-motor EV mode and both generator and traction motor propel the vehicle. The sprag clutch, which acts as an OWC, prevents the ICE from rotating in opposite direction when torque of generator is channelized from sun gear to ring gear. Technically, OWC serves the purpose of holding the carrier stationary so that torque from sun gear can be transmitted to ring gear. The powertrain dynamics exhibited by the dual-motor EV mode is modeled with a different set of kinematic relations as given by Eq.2.3.5. The schematic diagram of EV#2 mode is shown in Fig.2.13.(e).

$$J_{out_{eq}}\ddot{\theta}_{out} - \tau_{out} = \beta_1 * \{J_{gen_{eq}}\ddot{\theta}_{gen} - \tau_{gen}\} \quad (2.3.5a)$$

$$+ \beta_2 * \{J_{mot_{eq}}\ddot{\theta}_{mot} - \tau_{mot}\}$$

$$\ddot{\theta}_{gen} = -\beta_1 * \ddot{\theta}_{out} \quad (2.3.5b)$$

$$\ddot{\theta}_{mot} = -\beta_2 * \ddot{\theta}_{out} \quad (2.3.5c)$$

, where  $J_{out_{eq}}$ ,  $J_{ice_{eq}}$ ,  $J_{gen_{eq}}$ , and  $J_{mot_{eq}}$  are equivalent mass moment of inertia (MMI)s of transmission output, ICE, generator, and traction motor respectively. The equivalent MMI represents the combined MMI of the prime-mover and the associated gear of PGset.  $\beta_1$  and  $\beta_2$  is the ratio of number of teeth of ring gear to sun gear ( $\beta = \frac{R}{S}$ ) for 1<sup>st</sup> and 2<sup>nd</sup> PGset respectively.  $\tau$  and  $\ddot{\theta}$  represent torque supplied by and angular acceleration corresponding to each prime-mover.  $J_{out_{eq}}$  can be further disintegrated into rudimentary elements as given in Eq.(2.3.6)

$$\begin{aligned} J_{out_{eq}} &= J_{out_{at\ trans.\ output}} + J_{ring} \\ &= \left\{ \frac{J_{veh.}}{(FDR)^2} + J_{differential} \right\} + J_{ring} \end{aligned} \quad (2.3.6)$$

Considering straight road without any curves and absence of any lateral acceleration,  $\tau_{out}$  at transmission output is derived from vehicle loads and vehicle dynamics, as shown in Eq.2.3.7.

$$\begin{aligned} \tau_{out} &= \frac{\left( mg * \sin(\vartheta) + mg * \cos(\vartheta)(\varphi_1 * V + \varphi_2) \right) * r_{wheel}}{FDR} \\ &+ \frac{\left( mg * \left( \frac{\partial V}{\partial t} \right) + \frac{1}{2} C_d * \rho A_{front} V^2 + \tau_{brake} \right) * r_{wheel}}{FDR} \\ &+ \tau_{diff. loss} + \tau_{mesh loss} + \tau_{spin. loss} \end{aligned} \quad (2.3.7)$$

### 2.3.7 Pursuit of An Advanced Hybrid Architecture

Although powertrain electrification is undisputedly a disruptive technology, customers have not fully embraced the EV or HEVs as expected by the OEMs. Literature

clearly manifests the mix perception of the potential buyers of EV&HEVs across Europe [43, 44], U.S. [45], and China [46] towards this paradigm shift in sustainable mobility. While factors like range anxiety, longer charging time, and insufficient infrastructure aggravate buyer's perceived reliability towards full EV&HEV [45] and eventually have been deterring innumerable consumers away from buying EV&HEV [44], monetary and non-monetary incentives from governments have been registered as a positive factor that can ameliorate perfunctory behavior of the prospective customers of EV&HEV. However, it will be a dishonor to those buyers who are really concerned about climate change and have shown their interest to EV&HEV as their responsibility towards environment [47]. Nevertheless, to attract prospective buyers, the technology of powertrain electrification has to be more enticing, more robust, and more smart than its predecessor through a series of research collaboration between OEMs and governments.

Authors in [43] have proffered that combined effect of multiple factors (range anxiety, charging time, infrastructure, reliability, initial price, incentives, environmental concern, social reputation) govern the intention of buying an EV&HEV in a complicated way and that model needs to be understood by the OEMs and governments before investing colossal amount of fund in state-of-the-art research.

If the focus is narrowed down onto the hybrid-ePTs, which are not encumbered with range anxiety, initial higher price, and higher charging time, it can be seen that all the benefits of hybrid-ePTs have not been still squeezed out through research.

Toyota Prius is the most successful HEV till now since its introduction in Japan in 1997 and in North America in 2000. The Toyota Prius set an exemplary record of holding almost 50% and 70% market share of the HEVs in the U.S. in 2009 [48] and

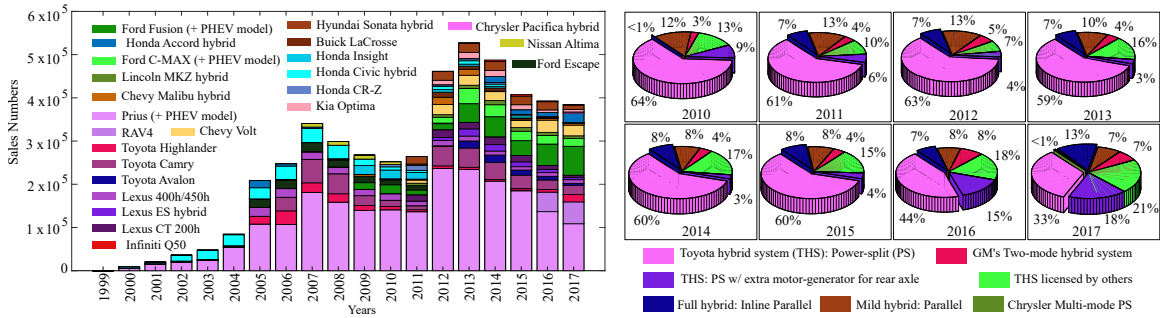


Figure 2.14: HEV sales trend in U.S. from 1999 to 2017.

in 2012 [5, 49] respectively.

Fig.1.1 reiterates the dominance of Toyota Prius and other Toyota HEVs in U.S. market since the penetration of electrified powertrain into automotive market. Although there are many factors govern the success or defeat of any HEV in terms of sales figure, sticker value of fuel-economy, drivability (ease of operation and driving comfort), and torque & power rating of the HEV, play a major role in its success in the market. Undisputedly, these factors primarily depend on the control strategy of the EMS and drivetrain architecture of the HEV. As we have compiled a trend of cumulative application of all commercial drivetrain architectures over the period of 2011 to 2017 in the Fig.1.2, it corroborates the success rate of second and third generation of THS. Simple operation and effective control of THS with two PG-sets have been so successful that other OEMs like Ford, Nissan licensed it and used in their HEVs with a few modifications. In the meantime, the automotive industry has experienced a resurgence of multiple ePT modes, introduced by GM in 2007, with two PG-sets and multiple clutches [6]. However, Fig.1.2 and authors in [8] reveal that GM's "two-mode hybrid system" could not gain decent popularity even after complicated topology with five operating modes (i.e., two all-electric modes and three hybrid-electric modes) and its prosperous attributes. On the other hand, the sales trend of Prius given in Fig.1.1

highly supports the fact how Toyota resisted the plummeting trend of Prius sales and recaptured the market by introducing third generation THS which was technologically more capable and comprised of two PG-sets instead of one PG-set that its predecessors carries. However, third and current generation of THS does not employ any clutch to facilitate multi-mode operation. The stark contrast between the popularity of aforementioned two drivetrain configurations with multiple PG-sets and clutches has ignited the research on advanced hybrid drivetrain configurations with multiple PG-sets and clutches. This sector of advanced power-split drivetrains has been acquiring attentions from a handful number of global research groups because the control policy of an EMS, whose prime objectives are to optimally assigning propelling load between ICE and EMG within their feasible limits and sustaining battery SOC, depends highly on the drivetrain topology [4].

## 2.4 Summary

Based on electrification factor, hybrid-ePTs are classified into micro, mild, parallel, series, power-split, and multi-mode ePTs. The power-split is the most successful configuration so far based on the sales figure. Both academia and OEMs are focusing on exploring new feasible configurations facilitating multiple electrified modes after observing this category's potential of catering various objectives. However, this category will incur more complexity in terms of both transaxle design and EMS articulation.



Table 2.1: Clutch and brake status in different modes in the predecessor “Two-mode Hybrid system” with Single PG-set

	Brake1	Clutch2	Clutch1
One-Motor All-electric	Closed	Open	Open
Two-Motor All-electric	Open	Close	Open
Series One-Motor Extended-Range	Close	Open	Close
Combined Two-Motor Extended-Range	Open	Close	Close

Table 2.2: Clutch and brake status in different modes in the successor “Two-mode Hybrid system” with double PG-sets

	B1	C1	One-Way Clutch
One-Motor EV	Closed	Open	Unloaded
Two-Motor EV	Closed	Open	Engaged
Low Extended Range	Closed	Open	Freewheeling
High Extended Range	Open	Closed	Freewheeling
Fixed Ratio Extended Range	Closed	Closed	Freewheeling

## Chapter 3

# Energy Management Systems: Fundamentals, Classification, and Utilitarian Aspect

## 3.1 Introduction

The prime objectives of transition from ICE based powertrain to ePT is reduction of both fuel consumption and tailpipe emission. Such objectives are undoubtedly achievable if energy generation of all the energy sources is optimally managed. Hence, the complexity incurred due to electrification will be well-justified if all the energy sources are optimally controlled.

A dedicated control system, known as EMS, has become integral part of the overall vehicle control module to carry out the above mentioned task. Theoretically, optimal control of more than one energy sources at all time-steps, given that driver's power request is satisfied, will engender a nonlinear constrained optimization problem. Global and local optimal solutions to such optimization problem are easily achievable with the help of many able control strategies and competent mathematical tools. Although the optimization is computationally intensive, they can be easily tackled with contemporary laboratory based digital processors. But the on-board microprocessor, dedicated for the EMS, is too delicate to handle computational burden of such optimization. Hence, researchers have been articulating robust and computationally mild algorithms for the utilitarian EMS of ePTs. The utilitarian EMS should conform to the memory constraint of on-board microprocessor and should yield optimal control decisions in real-time.

Most of the reported review articles in literature have dealt with theoretical aspect of different strategies solving energy management problem in electrified vehicles. The authors in [50] gave an overview about theoretical background of the basic control strategies to solve energy management problem corresponding to different hybrid architectures. A comparative study was conducted on four major theoretical energy

management strategies with corresponding simulation results [20]. Authors in [51] and [52] have presented a systematic classification followed by a detailed overview of contemporary energy management strategies. The authors in [53] reviewed different power management systems proffering modifications of rule-based control strategies to achieve near-optimal power sharing among the energy sources in real-time. The article highlights the importance of rule-based control strategies in real-time applications.

## **3.2 Energy Management System based on Powertrain Architecture**

The articulation of EMS relies strongly on architectures of ePT.

### **3.2.1 EMS for Micro and Mild Hybrid Architecture**

EMS in a micro hybrid vehicle is well-known as “auto stop-start” control strategy [54]. EMS for mild hybrid architecture retains auto stop-start control like micro hybrid but includes higher battery capacity and higher EMG power unlike micro hybrid. The mild hybrid architectures use 42V-48V battery system for either belt-driven starter-generator (BSG) [55] or integrated starter-generator (ISG) [56] or flywheel mounted electrical device (FMED). While most of the EMSs for micro and mild hybrid configurations are articulated with simple rule-based control strategy, optimization based control strategy can be also implemented to maximize the overall powertrain efficiency [57].

### 3.2.2 EMS for Strong Hybrid Architectures

The energy management problem for strong hybrid configurations is solved with various control strategies. The main philosophy of EMSs applied on series architecture is to operate the ICE at its best efficient region always due to the advantageous fact that ICE is totally decoupled from driveline for the series configuration [58, 59]. On a contrary, the ICE is not mechanically decoupled from the wheel either for parallel or series-parallel configurations. In parallel configuration, the ICE and EMG are coupled through torque coupling device, where torques from both the prime movers coalesce to serve driver's torque request. Many control strategies such as DP [60], extremum seeking (ES) [61], ECMS [62], adaptive-ECMS [63], [64], Adaptive-ECMS with fuzzy controller [65], robust proportional ECMS control [66], optimum fuzzy logic control [67], Bees algorithm [68], stochastic-optimal control [69], pareto control [70], Memetic algorithm [71], genetic algorithm (GA) [72], improved Particle swarm optimization (PSO) [73], fuzzy control with chaotic multi-objective GA [74], and calibration map-based optimal control [75] are employed to solve energy management problem for parallel architectures.

The series-parallel configuration can be achieved by employment of a clutch between series and parallel path of power flow. A single or dual planetary gear set is employed as a power-coupler before the final drive to unify the power flows coming from series and parallel path. Introduction of multi-mode operation with the help of planetary gear-set/s and clutch/es makes the task of the EMSs for series-parallel configuration even more challenging. To carry out such an challenging task, EMSs employ MPC [13, 76], nonlinear MPC through finite DP [77], SDP [78] artificial neural network (ANN)-based intelligent control [79], rule-based optimal control [80], velocity prediction and

nonlinear MPC [81], [82, 83] real-time near optimal control [84].

The significance of this section is to portray the variation in control strategy of EMS with respect to variation in architecture of the ePT. The number of energy sources that should operate optimally increases as the powertrain architecture become complex. The number of states as well as nonlinearity of the overall system dynamics increase as powertrain architectures evolves from simple to complex. Indeed, an evolution in the development of EMS has been taking place in parallel to the evolution of powertrain architecture. Such an evolution is depicted in the next section.

Powertrain architectures of PHEV are same as strong hybrid architectures with the only exception of enhanced all electric range (AER). Generally the EMSs, employed for PHEV, leverage the enhanced AER by operating the powertrain components either partially in charge depleting (CD) mode and then charge sustaining (CS) mode [85] or in blended mode [78], [86, 87] throughout the drive cycle.

### **3.3 Fundamentals, Evolution, and classification of energy management systems**

#### **3.3.1 Objectives of Energy Management Systems**

The predominant objectives of the EMSs are reducing dependence on fossil fuel and minimizing tailpipe emission of pollutants [2]. The EMS should satisfy the power demand from the vehicle and should constrain battery SOC within allowable limits. While copious number of researches have been focused individually either on reducing the fuel consumption [69, 88, 89] or minimizing tailpipe emission [75, 90], lion's share of EMSs has been articulated in order to satisfy both of the objectives together [91],

[61], [67, 68], [71], [74].

A handful of research works can be found dealing with non-conventional objectives such as minimization of degradation of battery health [92, 93] and drivability improvement [71]. Drivability issues, engendered from ICE start-up and gear-shifting, are mitigated using real-time implementable EMS [94]. Authors in [95] proposed a combined cost map considering weighted costs of fuel consumption, CO<sub>2</sub> emission, CO emission, hydro-carbon (HC) emission with respect to ICE operating points.

### 3.3.2 Working Principle of Energy Management Systems

Although different objectives are sufficed through a variety of control strategies and algorithms, the working principle of EMSs remain the same. Working as a supervisory controller, the EMS communicates with all the component-level controllers. The component-level controllers follow the commands from EMS to operate the EMGs, ICE and electric energy storage system (EESS) around their individual optimum operating points.

Since ICE is responsible for both fuel consumption and pollutant emission, control strategy of EMS should be competent enough to operate the ICE always within its optimum efficiency zone. The characteristic of ICE is such that the operating zone with maximum efficiency can only satisfy the objective of reducing fuel consumption but does not suffice the objective of minimization of tailpipe emission. Also, the constituents of pollutant family do not share the same optimum ICE operating zone as shown in 3.1.

Not only ICE, but EMGs and EESS should also be operated around their maximum efficiency zone to procrastinate capacity degradation of EESS. Optimal control is

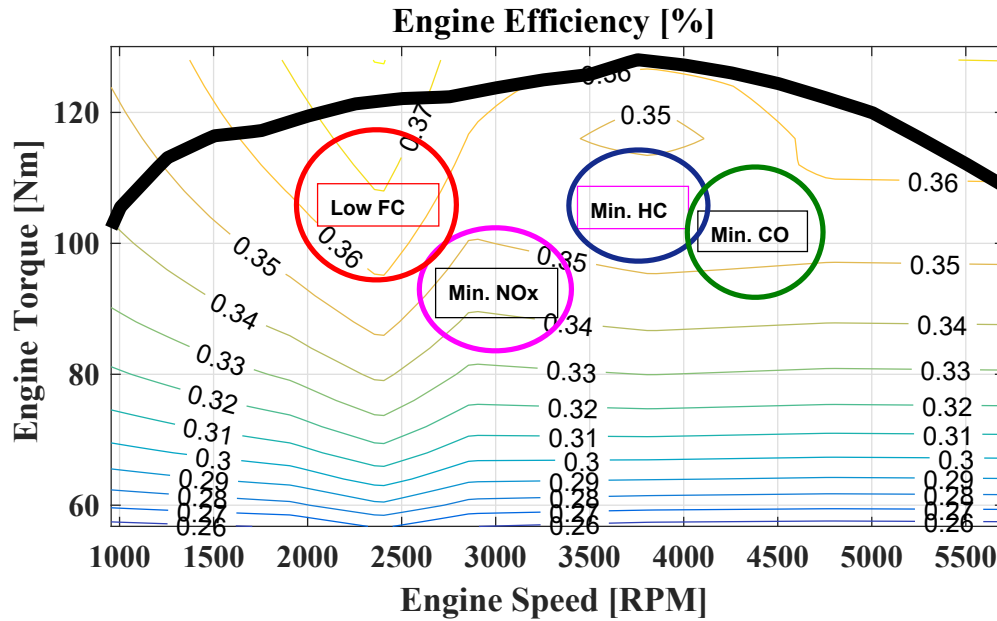


Figure 3.1: Representative optimum ICE operating zones with respect to different criteria for a typical ICE map.

the best sabre for the EMS to tackle such a convoluted responsibility. The EMS leverages the extra degree of freedom, in the form of electrification, by splitting the vehicle requested power with a variable quotient at each time-step and assigning the power-split between ICE and EMGs. As a benefit of electrification, EMSs of ePT can always operate ICE around its optimum points without compromising on the drivability objective. Assisted by the control strategy, the EMS emits a time-series profile of the power-split quotient and devolves it to the component level controllers.

### 3.3.3 Evolution of Energy Management Systems

A survey has been conducted to understand the trend and evolution of the EMSs. The survey includes literature satisfying the following criteria:

- Literature reported since 1993 up to 2018



- EMSs in the literature must deal with ePT with ICE
- Electrified powertrain should not include fuel cell
- Electrified powertrain should only include electric motors and should not include hydraulic or pneumatic motors

More than 250 research articles have been found satisfying the aforementioned criteria. The key discriminating characteristic of any EMS is its control strategy for energy management. 15 major control strategies such as deterministic rule-based, fuzzy rule-based, deterministic DP, QP, game theory (GT), GA and multi-objective GA, convex optimization, PSO, ANN-based, analytical optimization, PMP, different versions of ECMS, SDP, different versions of MPC, and RL-based control have been employed in different EMSs. Each of these strategies has their individual advantages and disadvantages and these are enumerated in 3.1.

Throughout the review of these research articles, it has been witnessed that scholars from different institutions and organizations have focused on different control strategies as per their preference around the same time-frame. Some group of scholars exploited a particular area of control strategy whereas, some other group explored different control strategies for energy management systems. Moreover, most of the aforementioned control strategies have their own mathematical foundation and are distinct from other strategies. Hence, it is very difficult to describe the evolution of control strategies, i.e. how a particular strategy evolved from its predecessor. However, there is an explicit evolution in the objectives of the EMS throughout two and a half decades as depicted in 3.2.

Initially, EMSs were concerned with splitting power request between ICE and EMG

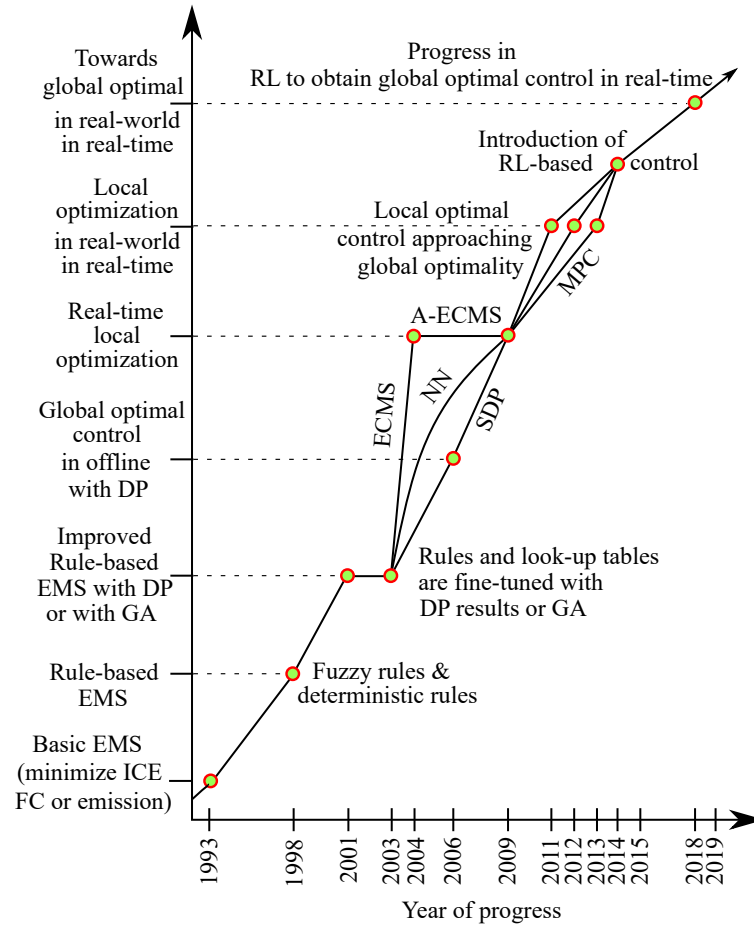


Figure 3.2: Tentative evolution of EMSs since 1993 till 2018.

so that emission from ICE or fuel consumption of ICE can be reduced [96]. Those control strategies were substantially rule-based and were constructed from dynamometer test and road test data [97]. These rules were never articulated focusing on optimal power-split between ICE and EMG until the assistance of DP was sought to extract near optimal rules for EMSs [98] or GA was employed to tune parameters of rule-based control [99]. With the advent of DP, optimal performance of ePT architectures is benchmarked with offline global optimization [100]. However, power-split control obtained from DP was impeded with prior knowledge of drive cycle and extensive

computational load and hence could not be implemented as a real-time optimal control.

Researches had been going on during that contemporary time in three distinct directions to circumvent the drawbacks of DP and to implement optimal control in real-time. Employment of machine learning through ANN [101], instantaneous optimal control via PMP and ECMS [102], and SDP [10] proved to be promising in solving the energy management problem in real-time with near-optimal solution. As time progressed EMSs approached towards optimal solution through advancement in these three directions. EMSs became more concerned in yielding optimal control for real-world driving situations with improved SDP [103], Adaptive-ECMS [95], improved machine learning framework [79], and with the advent of MPC [13]. It is noteworthy to mention that EMSs were also evolving in obtaining optimal control for satisfying multiple objectives together rather than satisfying single objective [74], [104]. The development of EMSs gained a giant leap towards attaining global optimal control in real-time for any real-world driving scenarios with the introduction of RL as the control strategy [105].

Although finding an explicit evolution of the control strategies is a daunting task, an approximate trend of popularity for each of the major control strategies can be portrayed. The trend of popularity of a control strategy is articulated with the year of first application, annual number of application, and the year of last application. Moreover, it is completely justified to present such a trend in the context of this article and hence depicted in 3.3.

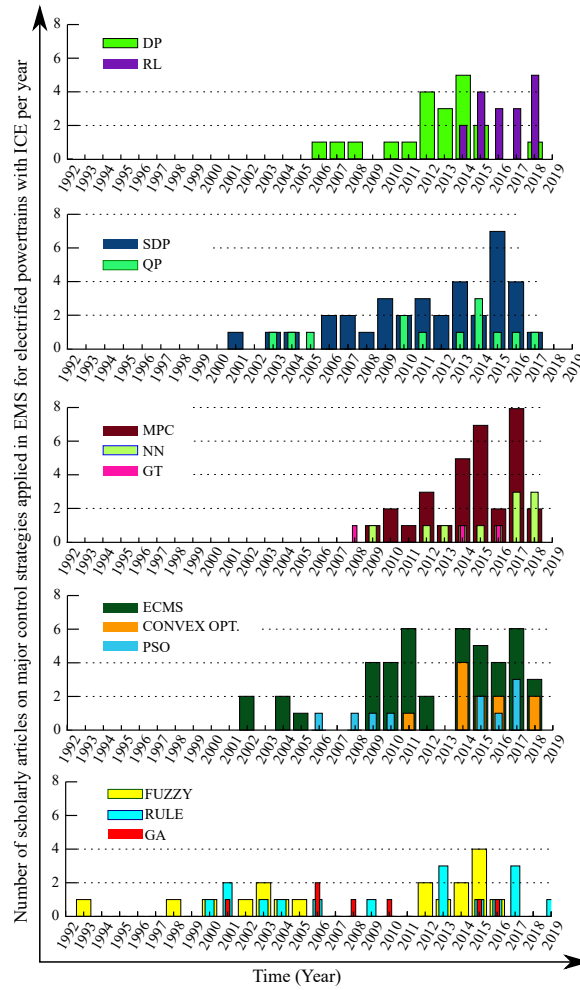


Figure 3.3: Trend of popularity of major control strategies used in EMSs

### 3.3.4 Classification of Energy Management Systems

Depending upon how the utilitarian EMSs are articulated to accomplish the assigned objectives, a novel classification is presented in this article. Utilitarian EMSs in real vehicle must execute in real-time. As far as EMSs in simulation environment are concerned, they are classified into offline and online categories. Although the online EMSs can execute in CPU-time within the simulation, they may not be apposite for real-time application. Hence, the full-fledged simulation-based online EMSs must go

through a simplification process before being encoded on the hardware of ECU. Such simplified version of EMSs are classified into:

- Premeditated energy management systems
- Casual energy management systems
- Blended energy management system

The classification is done based on the procedure followed by the utilitarian EMS to generate control decisions at each time-step. Each of these classes will be discussed elaborately in the subsequent three sections. Details of each class will include theoretical background, implementation in simulation environment, and steps of simplification before hardware implementation.

## **3.4 Premeditated Energy Management Systems**

### **3.4.1 Introduction**

These EMSs are comprised of precalculated control policy. The control policy is fixed and it is not altered until the EMS is reprogrammed via human intervention. Since the control policy is premeditated, EMS makes the control decisions almost instantly based on the values of inputs. This class of EMS performs only a few computations to make each control decision. Hence, this class of EMSs is not encumbered with computational complexity. The primary elements of this class of EMSs are input states, output control decisions, and a stationary control policy, as depicted in Figure 3.4. Before going into detailed steps of articulation, three questions should be answered chronologically.

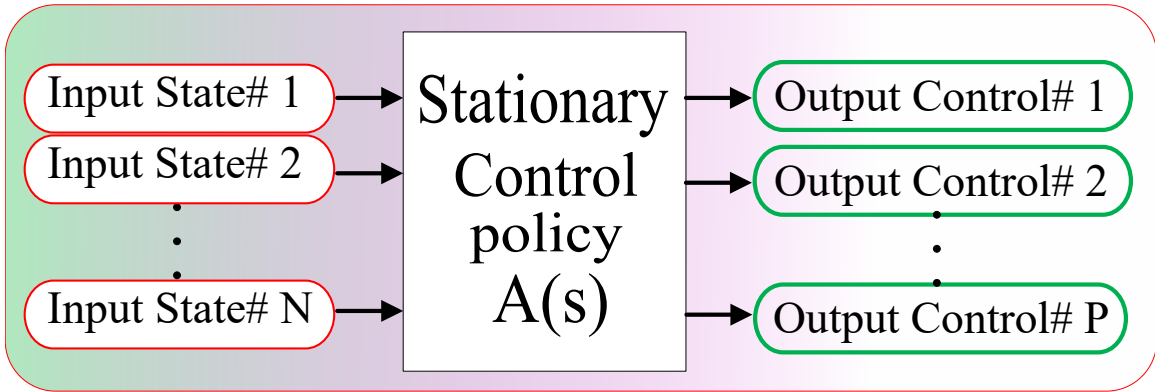


Figure 3.4: Primary components of premeditated EMS

- What are the forms of premeditated control policy?
- Which are the possible input states for the premeditated control policy?
- Are the control decisions optimal or near-optimal during real-life driving?

### Forms of Premeditated Control Policy

The premeditated control policy can be formed as a group of a few deterministic rules or as a group of a few fuzzy rules or as a look-up table. The policy with deterministic rules is comprised of certain IF-THEN criteria like the following:

- if state#1,#2,...,#N satisfy  $x^{th}, y^{th}, \dots, p^{th}$  conditions respectively, trigger  $A_k$  control action, where  $A_k \in \{A\}$
- if state#1,#2,...,#N satisfy  $r^{th}, y^{th}, \dots, z^{th}$  conditions respectively, trigger  $A_m$  control action, where  $A_m \in \{A\}$

A quintessence deterministic rule-based premeditated control strategy is shown in 3.5 as an example.

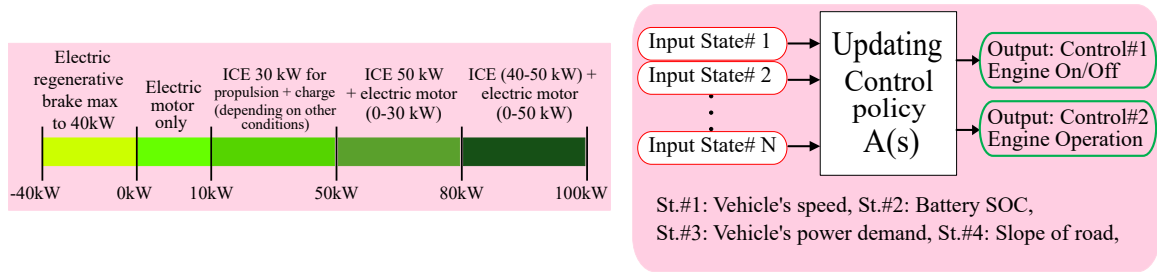


Figure 3.5: Example of deterministic rule-based control policy in premeditated EMS

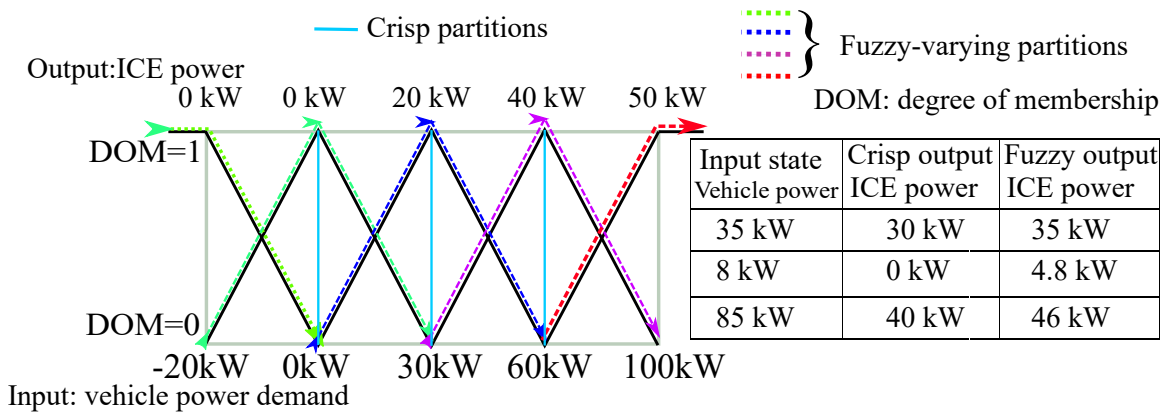


Figure 3.6: Example of fuzzy rule-based control policy in premeditated EMS

The premeditated control policy with fuzzy rules also relies on IF-THEN criteria. However, fuzzy rule-based control policy uses gradually varying partitions to define the conditions unlike the crisp partitions in the deterministic case. An example of fuzzy rule-based premeditated control policy is shown in 3.6. The third and most common form of premeditated control policy is a mapping of control decisions with respect to input states in the form of look-up table. Size of the dimension of the look-up table depends on the number of input states. Typical examples 1-D, 2-D, and 3-D look-up tables are shown in 3.7.

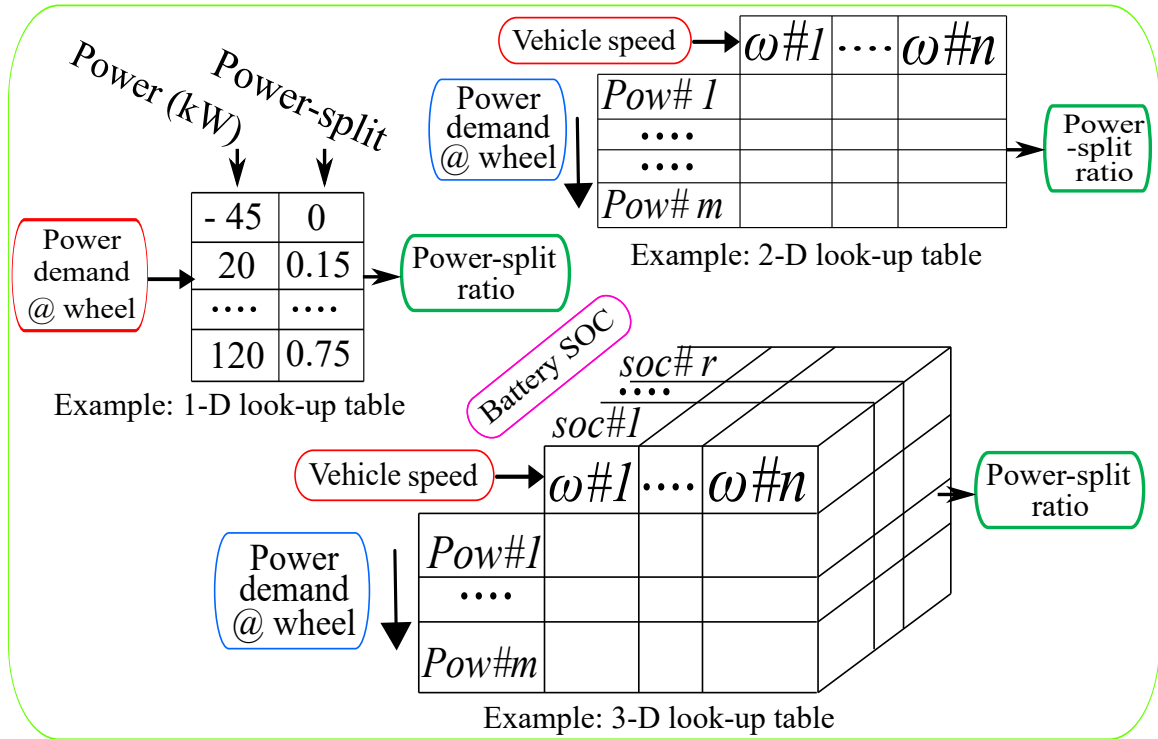


Figure 3.7: Typical look-up table based premeditated EMS

### Input states and Control Outputs of Premeditated Control Policy

Input states are the element, based on which control decision is made by the premeditated EMS at every time-steps. Torque or power demand at the wheel, vehicle speed, and gear-state of transmission are commonly used as input states to characterise the driving scenario. Battery SOC has been used as input state in almost every premeditated EMS to ensure charge sustainability.

Based on the values of input states, the premeditated control policy decides variables such as mode of operation, power-split ratio between ICE power and electric power, gear-shift command, ICE throttle command, and engine operating point command.



### **Optimality of Premeditated Control Policy under Real-World Driving**

Premeditated control policies are stationary and are articulated for a finite number of pre-known drive cycles. Hence, these control policies may or may not perform optimally in real-world driving scenarios. If the control policy is articulated focusing on specific drive cycles, the policy will perform optimally only when it encounter the same drive cycles during real-world driving. If the control policy is articulated focusing on probabilistic features of a constellation of driving situations, the policy will perform near-optimally even if it does not encounter same drive cycle during real-world driving. One advantage of articulating the premeditated policy based on driving situations rather than entire drive cycles is the performance of the control policy is still near-optimal for a real-world drive cycle which has not been optimized by the policy.

Due to operational ease, most of the original equipment manufacturers (OEMs) still rely on premeditated EMS for their electrified powertrains. However, the articulation of the premeditated control policy is a convoluted and rigorous task. The subsequent subsections will delineate various methods of articulating the premeditated policies.

#### **3.4.2 Detail Articulation Steps of Premeditated Control Policy**

The premeditated control policy can be articulated through two broad approaches. The first approach is post-processing of optimal control results obtained by solving the nonlinear constrained optimization problem. The second approach is gradual refinement of a stationary control policy through a systematic trial-and-error method known as a reinforcement learning.

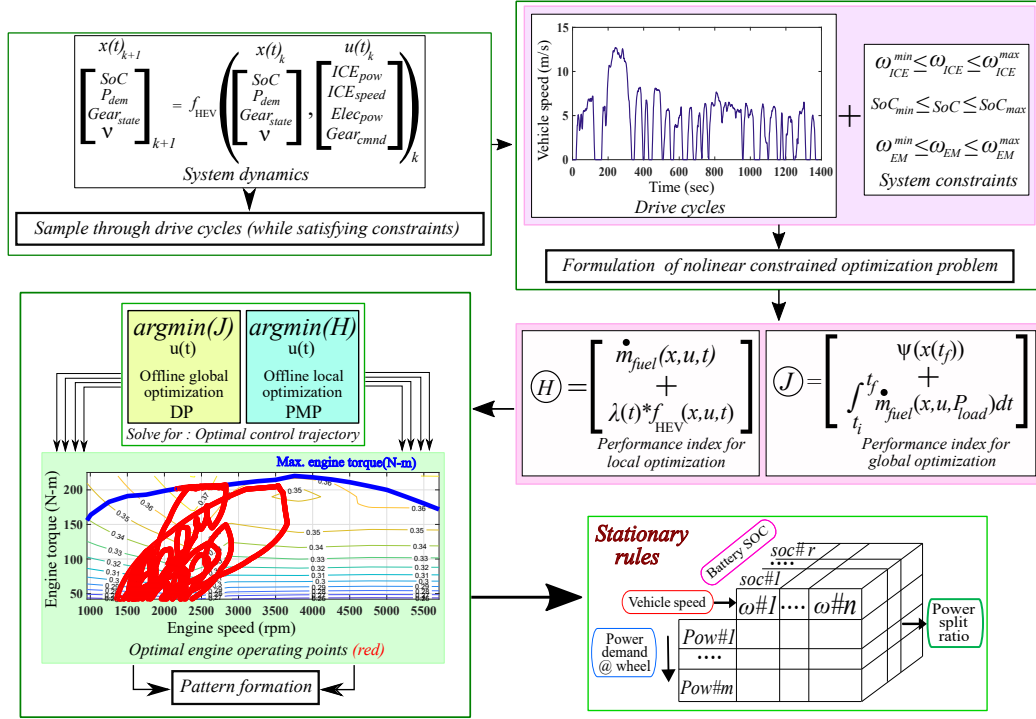


Figure 3.8: Basic articulation steps for premeditated EMS with post-processing

## Post-processing Based Articulation Steps

The post-processing based control policy articulation is comprised of optimal control trajectory generation and pattern generation, as depicted in the 3.8. Subsequent subsections will brief about the optimal control problem formulation and about two primary techniques to solve the nonlinear constrained optimization problem.

## Optimal Control: Problem Formulation

The optimal energy management problem for minimizing total fuel consumption is formulated as follows:

$$u^*(t) = \arg \min_{u(t)} \left[ \psi(x(t_f)) + \int_{t_i}^{t_f} \dot{m}_{fuel}(u, P_{load}, x, t) dt \right] \quad (3.4.1)$$

, where  $J = \left[ \psi(x(t_f)) + \int_{t_i}^{t_f} \dot{m}_{fuel}(u, P_{load}, x, t) dt \right]$  is the performance index,  $x(t)$  is the state vector of the system which is governed by the dynamics as follows:

$$\dot{x}(t) = f(u(t), P_{load}(t), x(t)) \quad (3.4.2)$$

, subjected to several state and control constraints such as:

$$\left. \begin{aligned} SOC^{min} &\leq SOC \leq SOC^{max} \\ \omega_{ICE}^{min} &\leq \omega_{ICE} \leq \omega_{ICE}^{max}; T_{ICE}^{min} \leq T_{ICE} \leq T_{ICE}^{max} \\ \omega_{EMs}^{min} &\leq \omega_{EMs} \leq \omega_{EMs}^{max}; T_{EMs}^{min} \leq T_{EMs} \leq T_{EMs}^{max} \end{aligned} \right\} \quad (3.4.3)$$

$u(t)$  is the control action vector,  $P_{load}(t)$  is vehicle load in term of power demand. The first term of the performance index, i.e.  $\psi(x(t_f))$ , represents a penalty due to deviation from end-SOC constraint. The second term represents fuel consumption cost over the full drive cycle which runs from initial time ( $t_i$ ) to final time ( $t_f$ ). Although the performance index varies for different combination of costs such as fuel cost, emission costs, structure of the minimization problem remain the same as in (3.4.1), (3.4.2), and (3.4.3).

Mathematical history refers two main solution strategies to solve such constrained optimization problem. DP [106] was proffered by Richard Ernest Bellman, an American applied mathematician, in 1957 and Pontryagin's maximum principle (PMP) [107] was proffered by Lev Semyonovich Pontryagin, a Soviet mathematician, in 1962.

## Optimal Control Trajectory Generation with Dynamic Programming

DP is a mighty numerical tool used as a solution technique for various domains such as optimal control problem, planning, and Markov decision process [106], [25]. DP is favorably applied for finding optimal solutions of complex problems which strictly can be disintegrated into sub-problems in discrete time step and each of the sub-problems has an optimal solution. Each of the sub-problems is associated with an instantaneous cost.  $v(s)$  is the cumulative cost obtained at the end of the drive cycle if the process initiates with state  $s$ .  $q(s, a)$  is the expected cumulative reward obtained at the end of the process if the process initiates with state  $s$  and action  $a$ .

Primary task of DP is to find minimum value functions  $v_*(s)$  and  $q_*(s, a)$  and the optimal control policy  $\pi_*(s|a)$  can be found from those minimum value functions.

## Articulation of Premeditated Control Policy from DP Results

Although DP is a powerful tool for solving non-convex optimal control problem, it is computationally intensive and cannot be implemented in real-time. Hence, a pattern is searched in the variation of optimal control trajectory with respect to certain states to avail DP-based control in real-time. If any particular pattern is found, it is converted into premeditated rules. But, if not found, a look-up table is articulated. Post-processing methods like least-square curve-fitting [108], optimization-based recalibration [109] are used to extract rule from optimal time-series profiles of control variables. Rules can also be deduced through finding a common pattern across the optimal results obtained for different driving situations [110, 111].

## Optimal Control Trajectory Generation with Pontryagin's Minimum Principle

With the advent of PMP, the constrained nonlinear optimization problem, shown in (3.4.1)-(3.4.3), can be solved analytically [17]. PMP is used in optimal control to determine the best plausible control action to take a dynamical system from one state to the next state.

Application of PMP transforms the constrained minimization of the performance index ( $J$ ) into minimization of the Hamiltonian  $H(x, u, t, \lambda(t))$  [107,112]. The necessary condition for obtaining a optimal trajectory of control action  $u^*(t)$  for all  $t \in [t_i, t_f]$  is as follows:

$$H(x, u, t, \lambda(t)) \geq H(x, u^*, t, \lambda^*(t)) \quad \forall u \quad (3.4.4)$$

,where the Hamiltonian is defined as:

$$H(x, u, t, \lambda(t)) = \dot{m}_{fuel}(u, x, t) + \lambda(t) * f(u, x, t) \quad (3.4.5)$$

with  $\lambda(t)$  is the vector of co-states. The first term of the Hamiltonian represents instantaneous fuel consumption and the second term represents equivalent electric energy consumption. State and co-state should satisfy the following conditions respectively:

$$\dot{\lambda}(t) = -\frac{\partial H(x, u, t, \lambda(t))}{\partial x} \quad (3.4.6)$$

$$\dot{x}(t) = \frac{\partial H(x, u, t, \lambda(t))}{\partial \lambda} \quad (3.4.7)$$

(3.4.4)-(3.4.7) have to be satisfied for the local optimality. The co-states always have

a physical meaning in optimal control problem. The co-state in energy management problem for electrified powertrain represents equivalence factor (EF) between electric energy consumption and fuel consumption.

### **Articulation of Premeditated Control Policy from PMP Results**

PMP generates instantaneous optimal solutions by minimization of the Hamiltonian at every time-step. Control policy obtained from PMP can be global optimal with certain reasonable assumptions [17]. However, the calculation burden of minimizing the Hamiltonian at each control step still may pose hindrances against implementation of PMP in real-time [89, 113]. The EF plays a key role in the optimality of the control policy obtained from PMP. Optimal value of EF is a function of time. However, optimal value of EF can be considered as constant with reasonable assumption [89]. In both constant and time variable cases, the main problem is that the value of EF depends on future drive cycle. Hence, it is mandatory to take an initial guess of the EF value for real-time implementation in utilitarian EMS.

The value of EF approaches towards its optimal value during real-world driving situations with the aid of linear [22] or nonlinear [114] feedback. Numerical PMP can be used offline to find optimal control policy for different discretized values of EF and the obtained policy can be stored in look-up table with respect to different discretized values of  $T_{demand}$ ,  $\omega_{wheel}$ , and EF [75, 89].

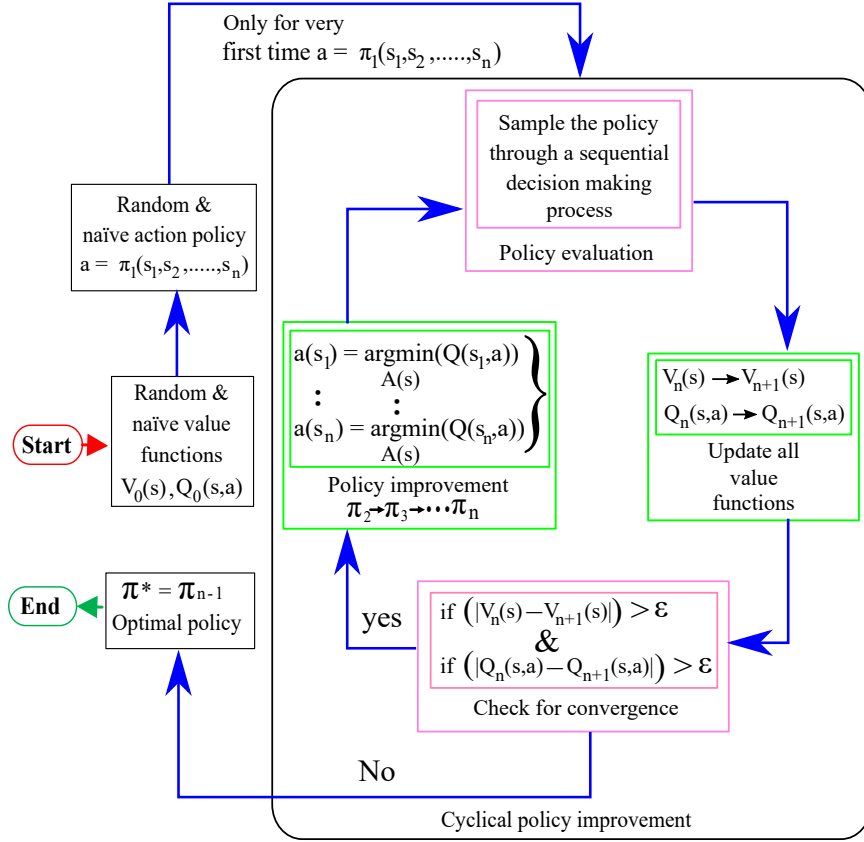


Figure 3.9: Cyclical repetition of policy evaluation and improvement

### Learning Based Articulation Steps

While post-processing based approach generates optimal control trajectory for different drive cycles one-by-one and articulates a common stationary policy for an ensemble of drive cycles, learning-based approach can articulate the common stationary policy for the whole ensemble through offline computation. This approach engenders a premeditated optimal policy through a series of progressive refinement of an initial naïve control policy, as shown in the 3.9. Each stage of refinement consists of policy evaluation and policy improvement.

The learning algorithm forecasts the expected long-term return of being in a certain

state (state value function,  $v(s)$ ) and expected long-term return of taking a certain control action at a certain state (action value function,  $q(s, a)$ ) in the policy evaluation step [25]. State of the environment is represented by “ $s$ ” and “ $a$ ” represents control action of the EMS controller. The forecast of  $v(s)$  and  $q(s, a)$  under policy  $\pi$  is done with following equation:

$$v_{\pi}(s_i) \doteq \mathbb{E}_{\pi}[Cost_{immediate} + \gamma * v_{\pi}(s_{t+1}|s_t = s_i)] \quad (3.4.8)$$

$$q_{\pi}(s_i, a_i) \doteq \mathbb{E}_{\pi}[Cost_{immediate} + \gamma * q_{\pi}(s_{t+1}|s_t = s_i, a_t = a_i)] \quad (3.4.9)$$

The action, which minimizes the action value function for every state, is chosen in the policy improvement step. The policy improvement step is executed with the following equation:

$$\chi^{(s_i)} \doteq \arg \min_{a(t)} \mathbb{E}[Cost_{immediate} + \gamma * v_{\pi}(s_{t+1}|s_t = s_i, a_t = a_i)] \quad (3.4.10)$$

Similar to DP, learning based approach finds minimum value functions  $v_*(s)$  and  $q_*(s, a)$  corresponding to all states and all state-action combinations respectively. But, it finds  $v_*(s)$  and  $q_*(s, a)$  for the probabilistic behavior of a constellation of driving situations instead of a single drive cycle [69].

Premeditated control policy obtained through offline learning-based approach can be directly used in production vehicles with diminutive manual tuning and re-calibration [103]. The derived control policy does not take much memory space if the state-space is two dimensional. As the number of state-space dimension increases, it becomes difficult to store high dimensional look-up table within ECU memory. Hence, size



reduction techniques such as, use of sparse matrix and use of coarse discretization, are implemented [69]. However, increase in dimension is beneficial for capturing real-world driving scenarios more accurately. Assistance of artificial neural network (ANN) has been sought to handle 4-D look-up table [115].

## 3.5 Casual Energy Management Systems

### 3.5.1 Introduction

The EMSs, which are neither dependent of precalculated control policies nor of the future knowledge of real-world drive cycle, are classified into casual EMS. Unlike pre-meditated EMSs, casual EMSs have to execute major calculations at each time-step to yield control decision. The casual EMS are expected to produce optimal or near-optimal control actions in real-time during real-world driving scenarios. Hence, the EMS should be able to perform a real-time optimization. The optimization algorithm should be simple enough to emit optimal or near-optimal control within control time-step. Also, the optimization algorithm of casual EMS must employ an efficient and adroit search method to find out the optimal control action in real-time. Following critical questions must be answered before developing a casual control policy for a utilitarian EMS:

- What type of optimization would be performed to emit control action?
- What search method would be employed to find the optimal control in real-time?
- What would be the inputs and outputs of the Casual EMS?

## **Type of Optimality of Casual EMS**

In general, local optimization is performed to obtain optimal control in real-time considering the contemporary computational capability of vehicle ECUs. However, with certain reasonable assumptions local optimal control can satisfy global optimality [17]. The most frequently used local-optimal control strategy is the evolved version of ECMS [22, 65]. The casual EMS with local optimization has to perform two tasks at each simulation time-step.

- Evaluate the performance index or cost for all admissible combinations of control variables.
- Find the values of control variables corresponding to the minimum performance index

## **Real-Time Search Method Employed in Casual EMS**

The aforementioned two tasks are accomplished by a search method during real-time implementation. There are three search methods available such as vector method, look-up table based method, and analytical method [116]. Vector and look-up table based methods can be implemented in real-time.

## **Inputs and Outputs of Casual EMS**

A typical casual EMS would require estimated value of vehicle's power demand  $P_{dem}$ , estimated EF  $\lambda$ , and battery SOC as inputs and it outputs power-split ratio between ICE and EMGs, mode of operation command for multi-mode transmission, gear command for automatic transmission, and engine start-stop command. Although the

calculation steps in the casual EMS are bit more complex than that of premeditated EMS, they will never let the EMS yield a control decision which is drastically far away from optimal decision. This is one major advantage of casual EMS over premeditated EMS.

### 3.5.2 Detail Articulation Steps of Casual EMS

Majority of casual EMSs are derived from ECMS. Although RL algorithm can theoretically output global optimal control policy in real-time, no utilitarian EMS, using RL algorithm, has not been reported in literature.

#### ECMS based Casual EMS

At each control time-step, an instantaneous optimization is performed to find local optimal control variables by minimizing the performance index. The performance index is known as Hamiltonian, given in (4.2.2). The Hamiltonian is evaluated for  $N_t$  number of iterations within a single time-step.  $N_t$  is the number of admissible values of control variable  $u_i$ . If there are more than one control variable,  $N_t$  will be the number of admissible combinations of  $u_i$ s. As stated before, vector and look-up table based methods are used to calculate Hamiltonian for all admissible combinations of  $u_i$ s.

Apart from the aforementioned tasks, casual EMS should estimate the value of EF. As stated in the last section, EF ( $\lambda(t)$ ) plays a major role in optimality of the control policy and it depends on future knowledge of drive cycle. In casual EMS, the value of ( $\lambda(t)$ ) should be updated in every time-step instead of keeping it constant throughout the drive cycle [22]. Proportional feedback control [66], nonlinear feedback control

[114], fuzzy rule-based control [65], adaptive control [63, 117] can be implemented to update  $(\lambda(t))$  in every time-step. The task of real-time update of EF can be devolved to the ECU dedicated for battery management system (BMS) to reduce the computational burden on the ECU dedicated for EMS [22].

### **Deterministic Dynamic Programming based Casual EMS**

Although deterministic dynamic programming (DDP) is beyond the thought of implementation in real-time EMS, DDP can be implemented with certain modifications, assumptions, and assistance of MPC. Reduction of the search grid, reduction of future window size, and approximation that DDP only chooses the most efficient ICE operating points can make DDP implementable in real-time [118].

## **3.6 Blended Energy Management Systems**

### **3.6.1 Introduction**

EMSs of this class are named as blended because they combine the benefits of both premeditated and casual EMSs. Blended EMSs alleviate the drawbacks of both premeditated and casual EMSs in terms of their real-time implementation. Premeditated EMSs are adroit in making control decisions since they need not to perform online optimization and their control decisions are globally optimal. However, premeditated EMS does not even guarantee near-optimality of the control decision when it encounters an unfamiliar driving situation during real-world driving.

On the other hand, casual EMSs output instantaneous optimal control decisions which

may or may not be globally optimal. Also, they must perform real-time optimization. Hence, the cost function used for the optimization should be either linear or quadratic for fast computation. However, casual EMS has one big advantage that it can guarantee instantaneous optimal control decision in any driving situation.

Real-time control policy obtained from utilitarian EMS can approach more towards global optimality if advantages of premeditated and casual EMS are combined to be complementary of each other. Some of the blended EMSs combine the advantages of premeditated and casual EMSs. Some EMSs which would not have been feasible to implement in real-time, have become feasible due to some precalculated rules and assumptions and henceforth included in the class of blended EMS. With the aid of some precalculated control rules, the blended EMS can make adroit control decisions online [119].

### 3.6.2 Detail Articulation Steps of Blended EMS

#### Reinforcement Learning Based Blended EMS

RL algorithm-based control strategy, discussed briefly in 3.4.2, cannot be implemented in real-time with computational power of contemporary automotive ECUs. Real-time implementation of this strategy is made feasible with little online computation and two pre-calculated separate look-up tables. One is for instantaneous cost ( $c(s_i, a_i)$ ) corresponding to all plausible state-action pair and another is for the value functions of plausible state-action pair ( $q_{\pi^*}(s_i, a_i)$ ) [120]. “ $s$ ”, “ $a$ ”, &  $\pi$  are defined in preceding section. Assisted by these two look-up tables, global optimal control is obtained in

real-time with (3.6.1), as follows:

$$a^*(s_i) \doteq \arg \min_{a(t)} \mathbb{E}[c(s_i, a_i) + \gamma * q_{\pi^*}(s_i, a_i)] \quad (3.6.1)$$

### **Adaptive ECMS Based Blended EMS**

As stated earlier, a feedback control is employed to regulate EF ( $\lambda(t)$ ) at each time-step along with finding local minimum value of Hamiltonian in ECMS based casual EMS. The EMS has to find both proper value of EF and local optimal control decision corresponding to minimum value of Hamiltonian. If one of the tasks is devolved to a look-up table, computational complexity will be reduced. Global optimal control policy corresponding to different EF values can be obtained from offline simulations over different driving situations and the policies can be stored in a look-up table. During real-time implementation, EF values can be calculated at each time-step via feedback control and global optimal control corresponding to the calculated EF value can be obtained from look-up table [75].

To expedite the computation of online local optimization process, use of pre-evaluated optimal equivalence factor is a smart option. Authors in [121] stored the optimal trajectories of equivalence factor with respect to different initial SOC values and different segments of drive cycle in the form of 2-D look-up table.

### **Model Predictive Control Based Blended EMS**

Model predictive control (MPC) can be implemented in a blended EMS with the assistance of a set of look-up tables and an online search algorithm [119]. Multi-parametric programming is used to articulate a stationary control policy in the form

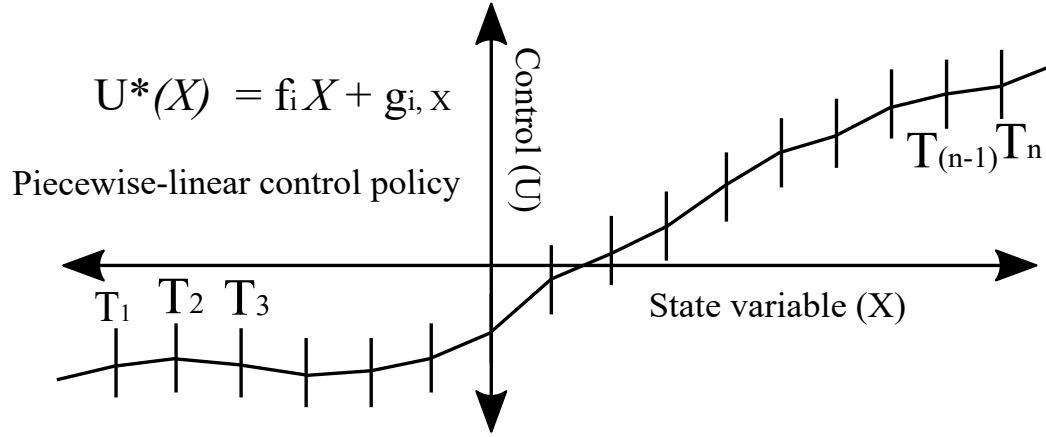


Figure 3.10: Example of control policy obtained from multi-parametric programming

of a continuous piece-wise-linear function of different state variables as shown in 3.10, where  $U_*(X) = f_i X + g_i; X \in T_i; i = 1, 2, \dots, N$ .  $U$  and  $X$  are control vector and state vector respectively and are shown in 3.10 where  $T_i$  is discretized domains of state.

In real-time, the EMS only has to search the correct domain in which the state resides at current sampling point. Several searching algorithm such as interval trees [122], ray-shooting approach [123] can be employed to accomplish the aforementioned task.

### 3.7 Development of Utilitarian Energy Management System

Since the paper mainly talks about utilitarian EMSs, significant steps throughout their journey from theoretical articulation to hardware implementation should be briefly narrated along with their detailed classification discussed in sections IV, V, and VI. The journey of a utilitarian EMS from theoretical foundation to hardware

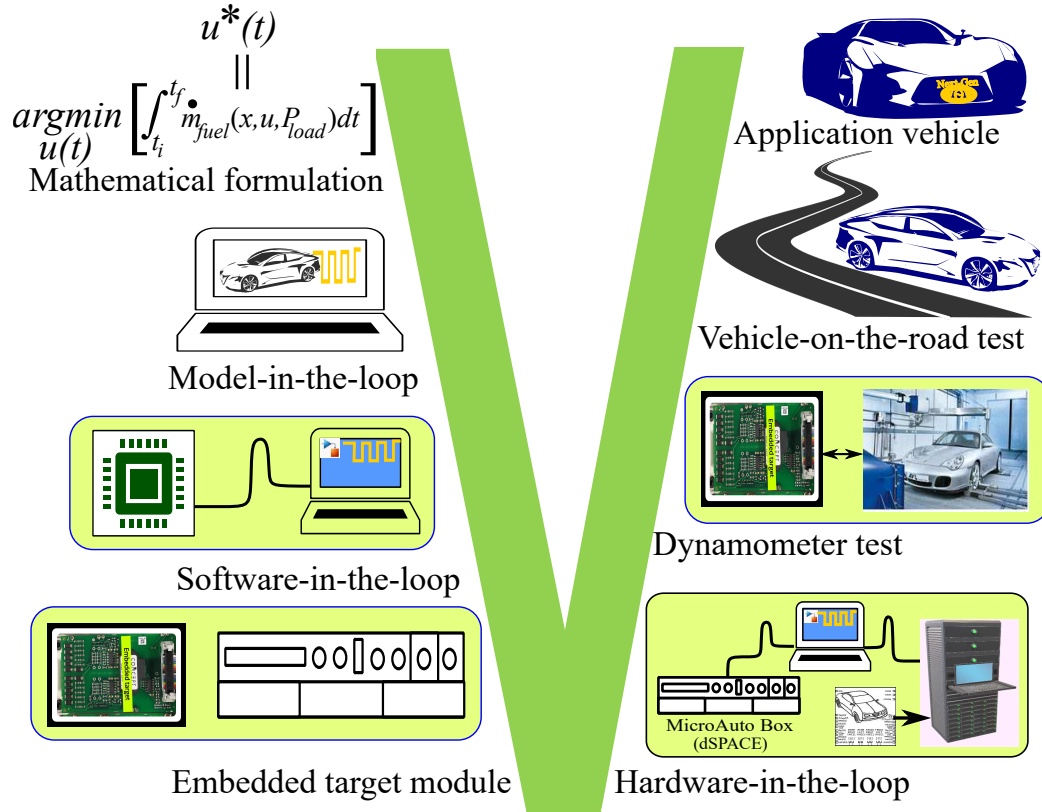


Figure 3.11: Chronological steps of EMS development

implementation is an iterative and a convoluted process. Hence, a systematic plan, as shown in 3.11, is followed to develop a utilitarian EMS.

The development process of a utilitarian EMS can be chronologically enumerated as follows:

- Model-in-the-loop simulation (MIL)
- Software-in-the-loop simulation (SIL)
- Hardware-in-the-loop simulation (HIL)
- Dynamometer test



- Vehicle-on-the-road test

### 3.7.1 MIL & SIL Simulation

MIL testing is done in a simulation environment where the plant model and the EMS controller are typically designed in MATLAB & Simulink® platform. MATLAB & Simulink® is ideal for designing system-level controllers like EMS.

Once the EMS performance is validated with model-in-the-loop (MIL) simulation, the control policy should vary in the real-time executable controller. This stage is known as software-in-the-loop (SIL) testing. The vehicle plant model can be retained as Simulink® based as it was in MIL. However, AMESim®, DYMOLA®, MapleSim®, and a few other platforms can be used for high-fidelity plant modeling. GM uses another process of executing software-in-the-loop (SIL) test, known as Virtual SIL (V-SIL), without controller hardware [124].

### 3.7.2 Hardware-In-The-Loop Simulation

Modern day automotive electronic control unit (ECU)s cannot be integrated in the vehicle without validating their performance through hardware-in-the-loop (HIL) simulation. In comparison to real-vehicle test, HIL is a proven cost-effective test procedure. Also, the repeatability of the same test scenarios gives an extra edge to HIL simulation over prototype or on-road test facilities. Key aspects of HIL simulation are discussed briefly as follows:

## **Communication System for HIL Simulation**

Controller area network (CAN) protocol is an efficient communication system which is indispensable to facilitate flawless and swift data exchange among so many ECUs in HIL simulation. CAN allows all the ECUs to communicate among themselves without encumbering the microprocessors. Qualities like low cost of data transmission, fault-tolerance, centralized error diagnosis, robustness, and provision of data prioritization make CAN so popular and ideal choice for electronic data transmission in automotive vehicles.

## **Development of Target Module Controller**

Auto-code generation feature of Simulink® and other software like MotoHawk Green Hills is used to convert control strategy into C-code [119]. The auto-generated C-code is flashed into microcontroller through MotoTune software [119]. MicroAuto-Box® from dSPACE® can be used for rapid control prototyping which is convenient for frequent alteration of control policy.

## **Development of Real-Time Executable of Plant**

The plant model of the powertrain, simulated in HIL test bench, should have enough fidelity to capture crucial dynamics of the powertrain components and should run in real-time. It is noteworthy to mention that HIL simulations are performed in discrete time with constant time-step. Hence, executable of plant model should be compatible with fixed-time solvers. Simulink® based plant models can be easily converted into executable C-code. However, Simulink® models are mean-value and look-up table based models which cannot capture intricate dynamics of powertrain components

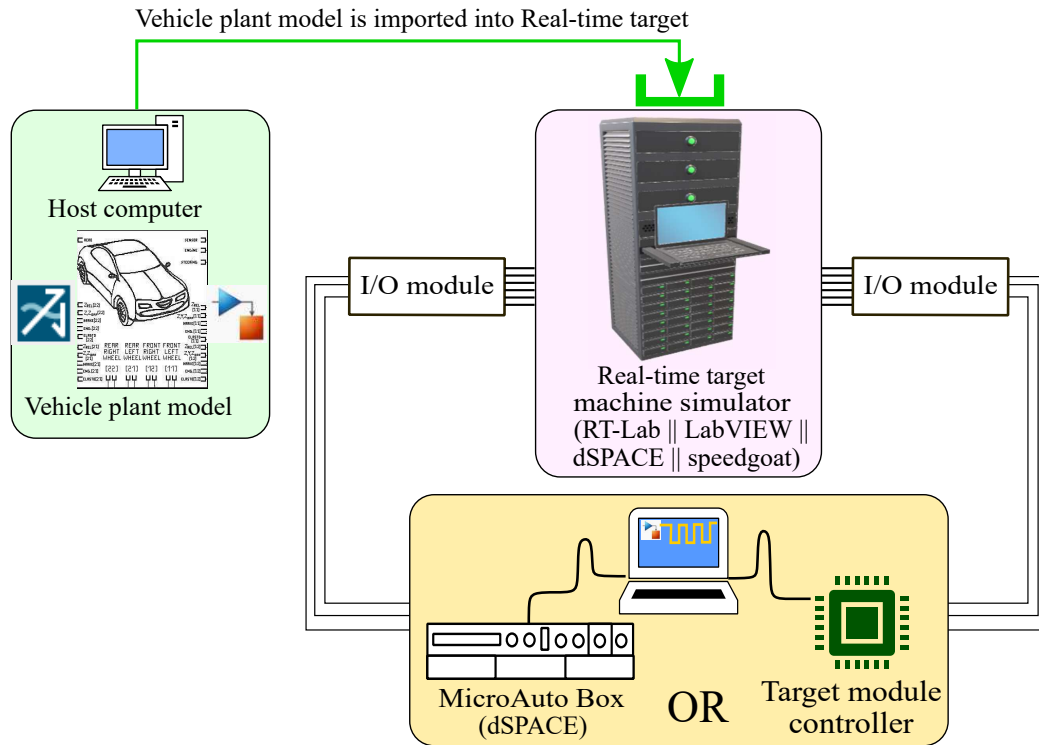


Figure 3.12: Schematic of hardware-in-the-loop simulation set-up

like differential equation based AMESim® models can do [125].

3.12 depicts a schematic of the HIL simulation test set-up. The high-fidelity plant model is developed in the top block and then imported into the hardware simulator, shown in the middle block. Simulator from different OEMs such as dSPACE®, RT-LAB®, LabVIEW®, and Speedgoat® can be used. As shown in the bottom block, the control strategy is articulated in MATLAB & Simulink® and then it is either exported to MicroAutoBox of dSPACE® or embedded onto a target module controller.

### 3.7.3 Dynamometer Test

Validation of the EMS controller on dynamometer (dyno) test-bench is the last level of controller development before the vehicle-on-the-road test. chassis dyno is used to validate the performance of the EMS controller with the full vehicle exposed to replica of different road load conditions [126]. In order to evaluate the performance of the controller convincingly, the same test-drive cycle should be repeated multiple times on the chassis dyno test [103, 126]. Chassis dyno test is the ideal phase where human participation should replace computerized PID controller-based driver.

### 3.7.4 On-road Vehicle Test and Calibration

After extensive validation through MIL, SIL, HIL, and dyno testing, finally the EMS controller is deployed in a production vehicle or a mule vehicle to validate performance of the controller on actual road conditions. After the road test, frequently it is found that the EMS control strategy requires a few modification because all the ECUs are communicating with each other in the real-world driving scenarios at this phase. It is not viable to do a major change of the control strategy at this stage because it is almost finalized for the production vehicle. Hence, the only inexpensive solution is to calibrate the controller. In other words, the control strategy is adjusted with respect to different driving condition to meet legislative norms and regulations from government in terms of fuel economy and emission and to meet customer's expectation in terms of performance and drivability.

## 3.8 Future Trends: Online Update of Energy Management Systems

So far in this article it has been described that with present computational capacity of ECU, real-time global optimal control cannot be implemented explicitly. Control policies in the previous three classes of EMSs approach towards global optimality in real-time in the following manner:

- stationary premeditated rules yield near-global optimal control in real-time.
- non-stationary but local optimal control seldom achieves global optimality.
- approximate global optimality of control is achieved with a combination of stationary premeditated rules and non-stationary local optimization.

The ultimate goal is to obtain a control policy which is non-stationary and globally optimal too. There are couple of plausible paths to reach the goal. Some of them are not feasible with contemporary capability of ECU but eventually will be feasible in future. Rest of the paths are feasible and researchers are working on it. This section will discuss all the paths briefly with their individual feasibility.

### 3.8.1 On-Vehicle Implementation of RL Algorithm in Real-Time

The RL algorithm optimizes the control policy for a constellation of driving situations which is represented by a transition probability matrix (TPM) of driving states. As stated in premeditated EMS, the control policy is optimized only once for a fixed

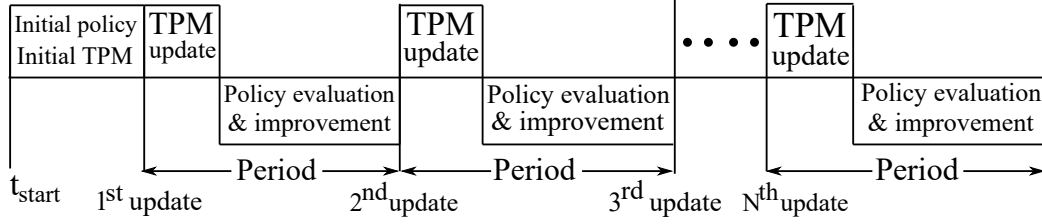


Figure 3.13: Concept of online update of control strategy with onboard computation

TPM. The drawback of premeditated EMS is that it may encounter a driving situation, which is not included in the TPM, during real-world driving. Hence, an ideal real-time global optimal control should update the TPM continuously in real-time and should yield global optimal control for the most recent TPM.

An EMS can be articulated with the explicit RL algorithm, an initial TPM, and an initial control policy. Then it can be deployed in the vehicle. The initial policy does not necessarily be globally optimal at the very beginning because the RL algorithm will eventually improve the policy to global optimality as it encounters more driving situations. As the vehicle is being driven on the road, the RL algorithm inside the EMS receives state signals from the sensors at each time-step. Then the EMS should update the TPM at every time-step and the RL algorithm should evaluate and improve the control policy also at each time-step. Therefore, these two tasks are performed continuously throughout every drive cycle. Although this strategy of implementing explicit RL algorithm seems to be enticing, it is not feasible with contemporary ECU's capability. The contemporary ECU can update the TPM at each time-step but it cannot evaluate and improve the control strategy at each time-step.

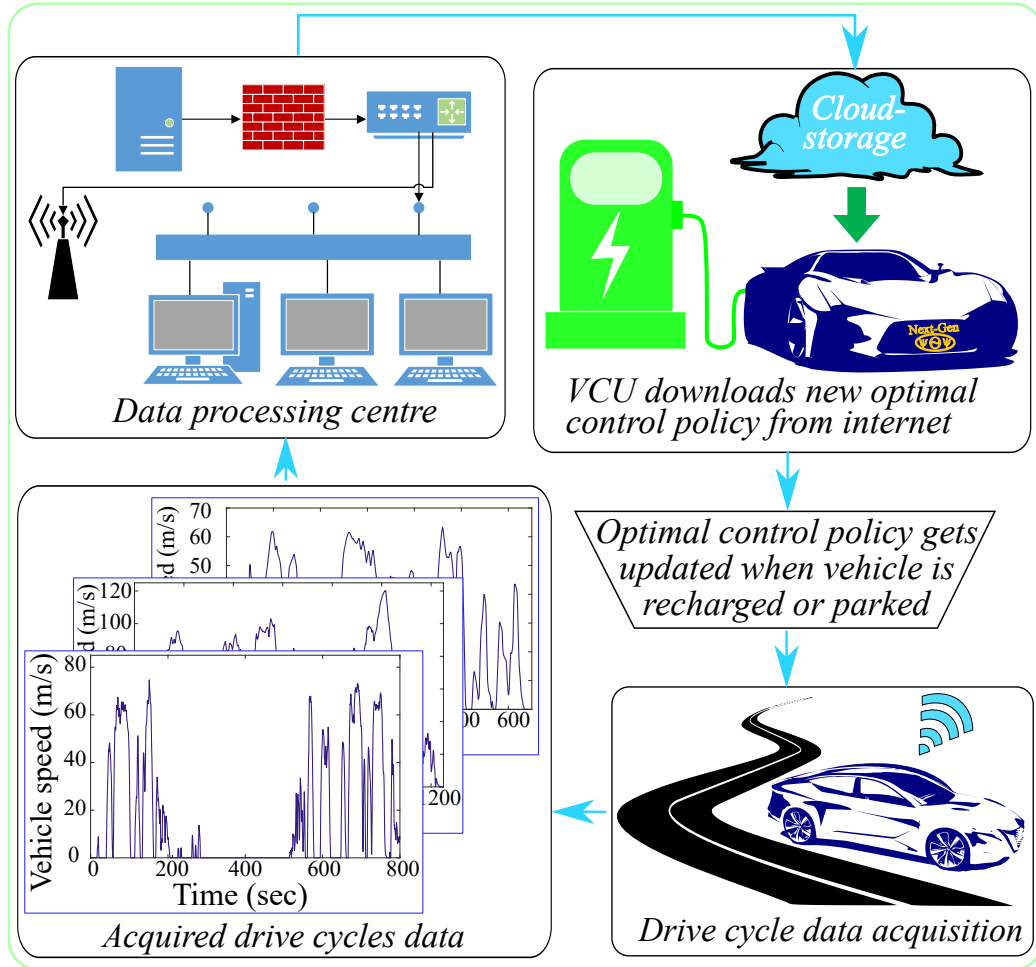


Figure 3.14: Concept of online update of control strategy for energy management system

### 3.8.2 Periodical Online Update of Control Policy with RL Algorithm

Looking at the computational limitation of ECU, two approaches can be prescribed. Since evaluation and improvement of the control policy are the computationally intensive part of the whole EMS, they are tackled by separate computer instead of an ECU. The computer can be on-board or off-board.

### **Periodical Update with On-Board Computer**

A separate advanced computer can be installed in the vehicle itself to replace the ECU dedicated for the EMS. The computer can continuously receive driving states from on-board sensors and can update the TPM periodically instead of at every time-step. Fraction of the period is used by the computer to update the TPM and rest of the period is utilized to evaluate and improve the control policy based on the most updated TPM as shown in 3.13.

### **Periodical Update with Off-Board Computer**

Another future trend is depicted in 3.14. The working principle is same as the previous one. But there is a difference in the execution. Instead of using on-board advanced computer, laboratory-based high performance computer is used here for periodical update of TPM, evaluation, and improvement of the control policy. Internet can be used as the mode of drive cycle data transmission from vehicle to research laboratory and updated control policy transmission from laboratory to vehicle.

## **3.9 Summary**

The total development process of a utilitarian EMS controller for electrified powertrain, starting from mathematical formulation of control policy to target module controller, is narrated in this paper. Before going deep into the main content, the typical characteristics of EMSs corresponding to major classes of powertrain architectures have been discussed. The paper also presents the tentative evolution curve of EMSs since 1993 till 2018. The curve will give an initial guidance to the future



researchers in this field and make them familiarize with the current scenario in EMS development at a glance.

It is shown that an optimal or near-optimal control policy for a utilitarian EMS can be articulated in the form of premeditated rules, instantaneous optimization, and combination of both. It is also delineated that the hardware constraints consistently play a major role in every phases of EMS development. The major focus here is to categorize the utilitarian EMSs based on their dependency on online computation or offline pre-computation or even both. In premeditated EMSs, it is shown that the optimal results are converted into robust stationary rules or look-up tables by finding a common pattern across the optimal control. The casual EMSs perform local optimization at each time-step with contemporary computational power of automotive ECUs. The novel classification strategy will eventually motivate future researchers to articulate EMS control strategies not only just for the sake of simulation-based performance but also for their real-time and real-world implementation. It is noteworthy to mention that both the trend of popularity and the comparison of control strategies have been presented here to help future and present researchers in selecting suitable control strategies for their utilitarian EMS. The five-stage validation process of a typical ECU is delineated systematically. The article entuses current and future researchers to brainstorm innovative solutions for a paradigm shift in the development of a utilitarian EMS.

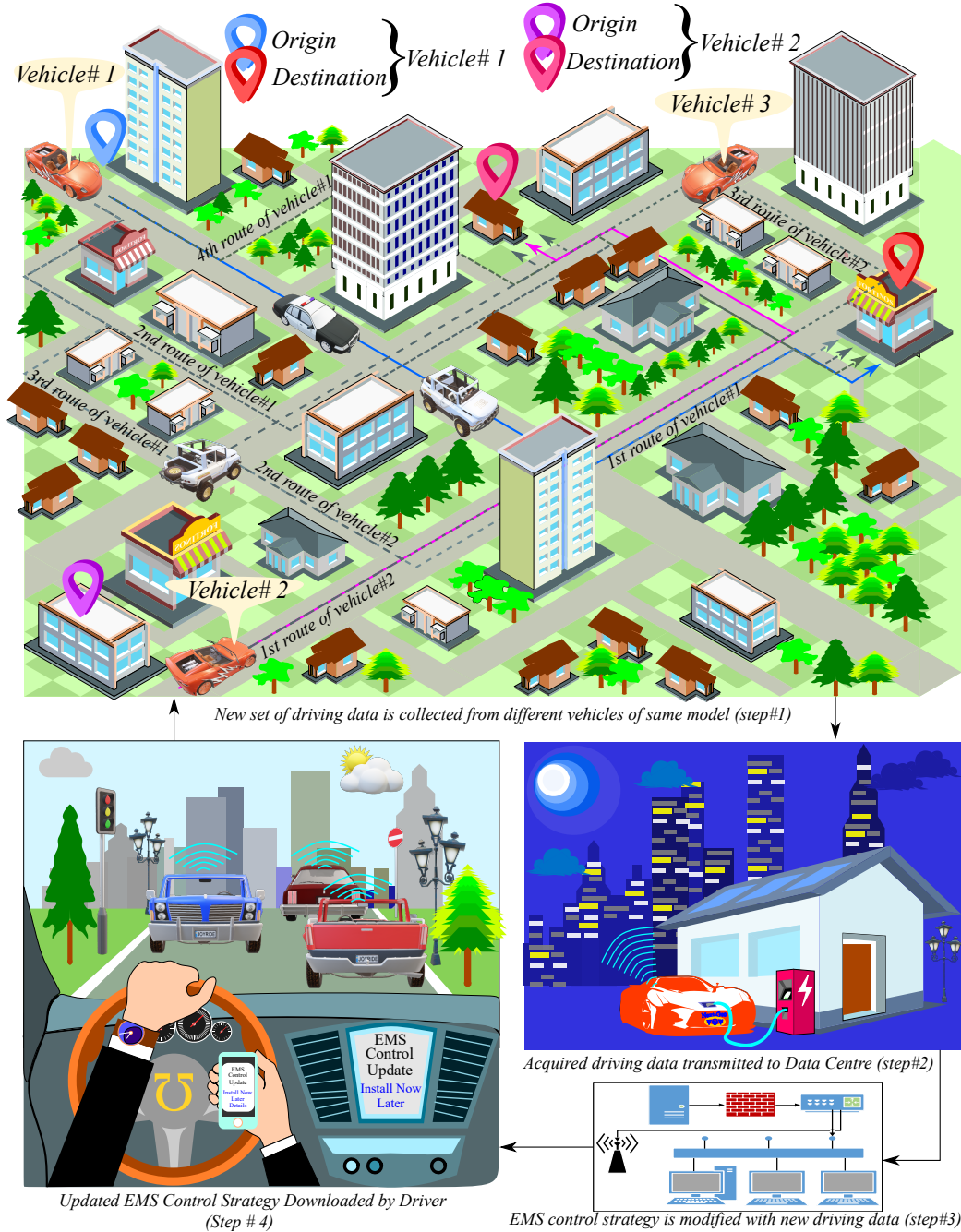


Figure 3.15: Concept of online update of control strategy for energy management system

Table 3.1: Comparison of control strategies used in EMSs

Control strategy	Advantages	Disadvantages
Rule	Simple to implement, consumes minimum ECU memory.	Depends heavily on intuition and experience.
Fuzzy	Easy to implement for nonlinear system with multiple objectives.	Difficult to tune parameters of membership function for global optimal control.
GA	Applicable to non-differentiable and discontinuous objective functions. Supports multiple objectives.	Prone to stuck in local optima. Convergence highly depends on initial population.
PSO	Less parameter tuning. Robust to initial population size.	Need extra memory to update velocity. Drive cycle dependence.
DP	Perfect global optimal solution.	Online implementation is infeasible. Heavy reliance on prior knowledge of future drive cycle.
Convex Optim.	Real-time optimization is feasible. Computationally efficient.	System dynamics must go through convexification. Global optimal control cannot be achieved.
QP	Subset of convex optimization. Real-time optimization. Commercial solvers are available.	Non-linearity is approximated with quadratic equation.
ANN	Less computational burden. Near-global optimal control for real-world driving.	Performance depends on lots of training data and prediction accuracy.
PMP/ECMS	Analytical solution of non-linear constrained optimization in real-time.	Local optimal doesn't guarantee global optimality.
SDP	Global optimal solution for a constellation of drive cycles.	Depends on the availability of Markov decision problem. Consumes a lot of time.
MPC	Inherent ability to tackle constraints on input, output, and states. Real-time optimization.	Depends on prediction accuracy. Seldom achieves a global optimal solution.
RL	Less memory usage. Continuous learning of the decision-maker. Robustness against unprecedented power demand.	No clear guidance in choosing immediate cost for multi-objective global optimization.

## Chapter 4

# Articulation of an Instantaneous Optimization based Control Strategy

## 4.1 Introduction

The objective of this dissertation is the pursuit of a near-optimal, intelligent, casual, and real-time control strategy for an EMS dealing with a multi-mode ePT. Before searching for such an ingenious EMS strategy, isn't it important to go through one of the most-accredited casual control strategy, i.e., equivalent consumption minimization strategy (ECMS). This chapter will elaborate the concept of ECMS with its implementation for two reasons. First, is to present the core concept of ECMS with its detailed implementation stages for a multi-mode ePT. After elaborating the basic and optimal ECMS, the author will propose a novel derivative of the basic ECMS with an ambition of achieving better performance than optimal ECMS. Second, the optimal ECMS has been institutionalized as a benchmark online optimization-based control strategy in literature. Hence, it is justified to compare any new real-time control strategy's performance with optimal ECMS's performance for proving the new one's effectiveness.

This chapter is comprised of detailed implementation of ECMS in discrete time-step making it apposite for embedded hardware, and a detailed implementation of a novel loss-minimization strategy having the same cornerstone as ECMS. This chapter will conclude with an avenue for a requirement of an advanced modeling of powertrain dynamics in a multi-mode architecture.

## 4.2 Equivalent Consumption Minimization Strategy: A Casual Incarnation of Pontryagin’s Minimization Problem

The most well-known casual EMS is referred as ECMS [15] which has been proven as the realization of application of PMP [16] in optimal control of energy management problem for HEPT. Both ECMS and PMP transforms global optimal control problem into instantaneous optimal control (IOC) problem by minimizing a Hamiltonian at each time-step, given that the co-state dynamics and state satisfy necessary and sufficient conditions respectively [16]. The general notation of the necessary and sufficient conditions are given as follows:

$$H(x, u, t, \lambda(t)) \geq H(x, u^*, t, \lambda^*(t)) \text{ (necessary)} \quad (4.2.1)$$

,where the Hamiltonian is defined as:

$$H(x, u, t, \lambda(t)) = \dot{m}_{fuel}(u, t) + \lambda(t) * f(u, x, t) \quad (4.2.2)$$

with  $x$ ,  $\lambda(t)$ , and  $u$  are the vectors of states, co-states, and control variables.  $u^*$  is the optimal control.  $\lambda^*(t)$  is the optimal time variation of the co-state.  $\dot{m}_{fuel}$  is mass flow rate of fuel consumption.

$$\begin{aligned} \dot{x}(t) &= \frac{\partial H(x, u, t, \lambda(t))}{\partial \lambda} \text{ (necessary condition)} \\ &= f(u, x, t) \end{aligned} \quad (4.2.3)$$

$$x(t_{initial}) = x(t_{final}) \text{ (necessary condition)} \quad (4.2.4)$$

$$\begin{aligned} \dot{\lambda}^*(t) &= -\frac{\partial H(x, u, t, \lambda(t))}{\partial x} \text{ (sufficient condition)} \\ &= -\frac{\partial \dot{m}_{fuel}(u, t)}{\partial x} - \lambda^* \frac{\partial f(x^*, u^*, t)}{\partial x} \end{aligned} \quad (4.2.5)$$

When the entire drive cycle information is not available apriori, both necessary conditions can be fulfilled with the co-state vector's adaptive variation. However, the sufficient condition required for global optimality cannot be satisfied without apriori knowledge of the drive cycle, and hence, the global optimality cannot be guaranteed. Although the co-state dynamics can be computed at every time-step with the available information about powertrain dynamics, global optimality comes with only a unique initial value of the co-state which must be found iteratively with the full drive cycle known in advance [127]. In [128], the authors have shown that PMP can give global optimal control without the fulfilling the sufficient condition regarding the co-state variation if the battery resistance and open circuit voltage vary infinitesimally small with SOC variation. This implies that when battery operates within the range between 50% – 70% SOC, then only there is no need to satisfy the sufficient condition regarding co-state variation for global optimality. This work posits that the necessary conditions can also be sufficient for global optimality if and only if the mass flow rate of fuel can be expressed as a convex function of battery power and also the SOC variation is a concave function of SOC. However, in both cases, either sufficient condition is required or not required to be fulfilled, the initial guess of the co-state vector is mandatory.

Even if the co-state variables are constant throughout the drive cycle with a few assumptions, their initial value should be iteratively found with advance knowledge of the entire drive cycle [128, 129]. If driving mission is known in advance, PMP or ECMS can yield optimal control of close proximity with DP [17]. It is noteworthy to mention that the co-state of PMP is nothing but a Lagrange multiplier and has been incarnated as the EF between chemical energy and electrical energy in ECMS [130]. Since advance knowledge of the entire driving mission is mandatory for calculating the unique initial value of the co-state variable(s), the only remaining approach is to update the variable(s) adaptively with an initial guess during online or real-time application when advance knowledge of the drive cycle is not available [18]. Consequently, such adaptive variation does not satisfy the sufficient condition related to co-state dynamics for global optimality.

It has become a common practice to update the EF or co-state variable(s) adaptively and a handful of adaptive methods have been proposed and reiterated in literature since 2005. The authors in [19] and [20] concurrently introduced the concept of A-ECMS where the EF was updated periodically based on past trend of vehicle speed and prediction of future driving mission. The authors in [130] employed an proportional-integral-derivative (PID) controller for the very first time to update EF at every time-step and many others followed similar approach either with PID [21, 131] or just proportional controller [22, 132]. Advanced adaptation rules have been articulated based on prediction of future driving condition [23, 133] and based on recognition of current driving pattern [24, 134, 135], [21, 131, 136], [22, 132], [23, 133]. Another simple but profoundly used approach is to map the optimal values of EF as a



function of battery SOC and vehicle's power demand and to use the map during real-time implementation. Infinite horizon SDP has been employed to derive a mapping of EF on the state variables for a statistical data set coming from different uncertain driving mission on a fixed route and this map can be used during real-time application [137]. Data-driven approaches, assisted by ANN, have performed well in real-time implementations when they had been trained beforehand with optimal EF trajectories for different real-world driving situations [138], [137, 139]. There are a few drawbacks of each of the aforementioned methods for real-time implementation of IOC, some of them are theoretical and some of them are from utilitarian perspective and they are enlisted as follows:

- PID or proportional controller-based adaptive methods are not completely independent of tuning parameters. Controller gains must be tuned to achieve charge sustainability.
- Prediction-based methods highly rely on the accuracy of the prediction of future speed profile or the recognition of driving pattern. Also, complicated ANN structures might be computationally too demanding for the utilitarian hardware.
- Map-based methods heavily rely on the offline computation and can only yield desired performance for selective drive cycle(s) and routes. Offline calculation needs to be repeated for every time the vehicle controller comes across a new route or driving mission. Moreover, memory constraint of contemporary hardware might hinder big map from getting flashed into real controllers.
- Albeit EF has a physical interpretation, it is difficult to find the initial guess of

EF which will satisfy charge sustainability criteria.

- Most of the literature presented the concept of instantaneous optimization but a few of them elucidate the potential challenges and roadblocks encountered during actual implementation and plausible ways to circumvent those roadblocks.

### 4.3 Implementation of Optimal ECMS

The previous section has presented a comprehensive review of state-of-the-art implementations of ECMS proposed in literature since 2001 until 2019 considering mostly unknown future driving situations. The reason for which so many researchers have been inclining towards ECMS since its introduction in 2001 is its operational ease even with the slight computational complexity of the instantaneous optimization, and its capability of yielding near-optimal control for any drive cycle. So, why are scholars still researching on different state-of-the-art derivatives of the basic ECMS? The answer is, the researchers are trying to circumvent the unavailability of future driving information during real-world and real-time driving, and they are striving for reducing the gap between the global optimal control and the ECMS-based near-global optimal control in unprecedented driving situations. As this dissertation will proceed towards culmination, the impact of a basic-ECMS on the targeted control strategy will be elucidated, especially for an EMS dedicated for a multi-mode ePT.

This section will delve into the detail implementation of optimal ECMS for a multi-mode ePT in discrete time-step. The entire simulation platform is developed in MATLAB-Simulink  $\text{\textcircled{R}}$ environment, as shown in Fig.4.1. The Simulink  $\text{\textcircled{R}}$ model is developed in many layers. Requirement and working principle of every layer will

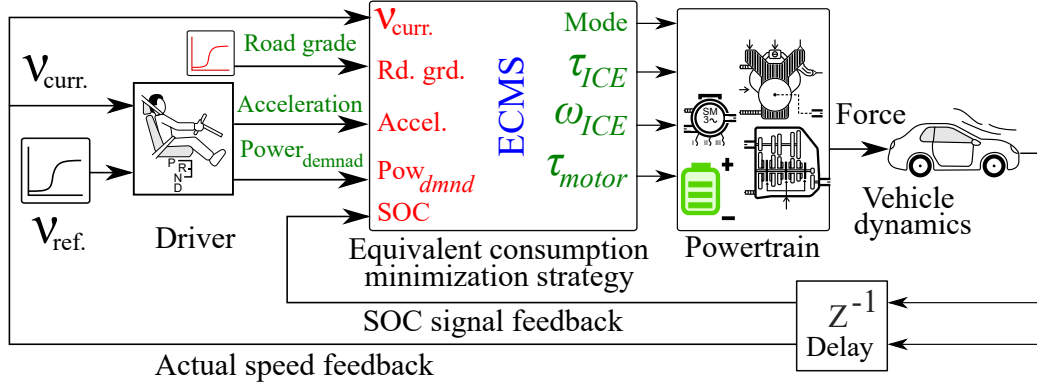


Figure 4.1: Overview of the simulation framework used for ECMS implementation

elaborated in the following sub-sections.

### 4.3.1 Overview layer

The overview layer depicts the interaction between three cardinal parts, i.e., the driver block, the ECMS controller, and the plant model of the simulation framework. 1-dimensional (1-D) longitudinal vehicular dynamics is considered as an integral part of the powertrain, and the plant model is comprised of the powertrain components and 1-D vehicular dynamics. Both ECMS controller block and driver block receive feedback signals from the plant model. The driver receives mainly the feedback signal of vehicle's speed whereas, the ECMS controller receives both feedback signals of vehicle's speed and battery SOC from plant model.

### 4.3.2 Driver block

The driver block emulates a human driver, and it compares the simulated speed and acceleration of the vehicle from the previous time-step with the reference speed and reference acceleration, respectively at the current time-step to adjust the accelerator

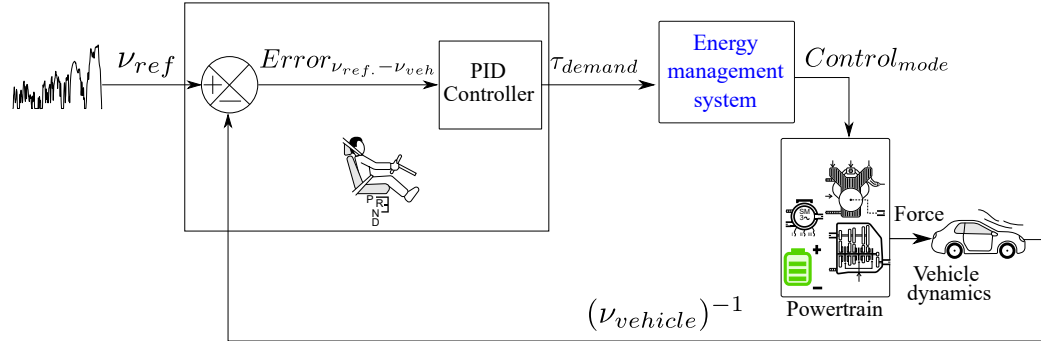


Figure 4.2: Schematic diagram of the driver block used in a model-in-the-loop simulation for vehicles

pedal position and brake pedal position. The most-accredited PID controller is employed as the cornerstone of the driver block. In this framework, the driver block spews vehicle’s torque demand instead of two separate output signals, i.e., change in accelerator pedal position and change in brake pedal position.

### 4.3.3 Plant model of Powertrain Components and Vehicular Dynamics

The midsize 2500 kg representative passenger vehicle, named as BaBaVan (kept in the memory of Iman), with a state-of-the-art multi-mode electrically variable transmission (e-VT) forward wheel drive (FWD) transaxle [42] is selected for this study. The schematic diagram of the multi-mode e-VT is shown in 2.13 in 2.3.6. Fig.4.3 helps to revisit to the three modes and their operating principle. The transaxle is comprised of an ICE, two EMGs, and a high voltage battery (HVB) pack. Since the objective of this study is to emulate a real-time interaction between the ECMS controller and the ePT, high-fidelity transmission models are used instead of steady-state models. The HVB, ICE, and EMG models are kept at low-fidelity. Inclusion of a high-fidelity

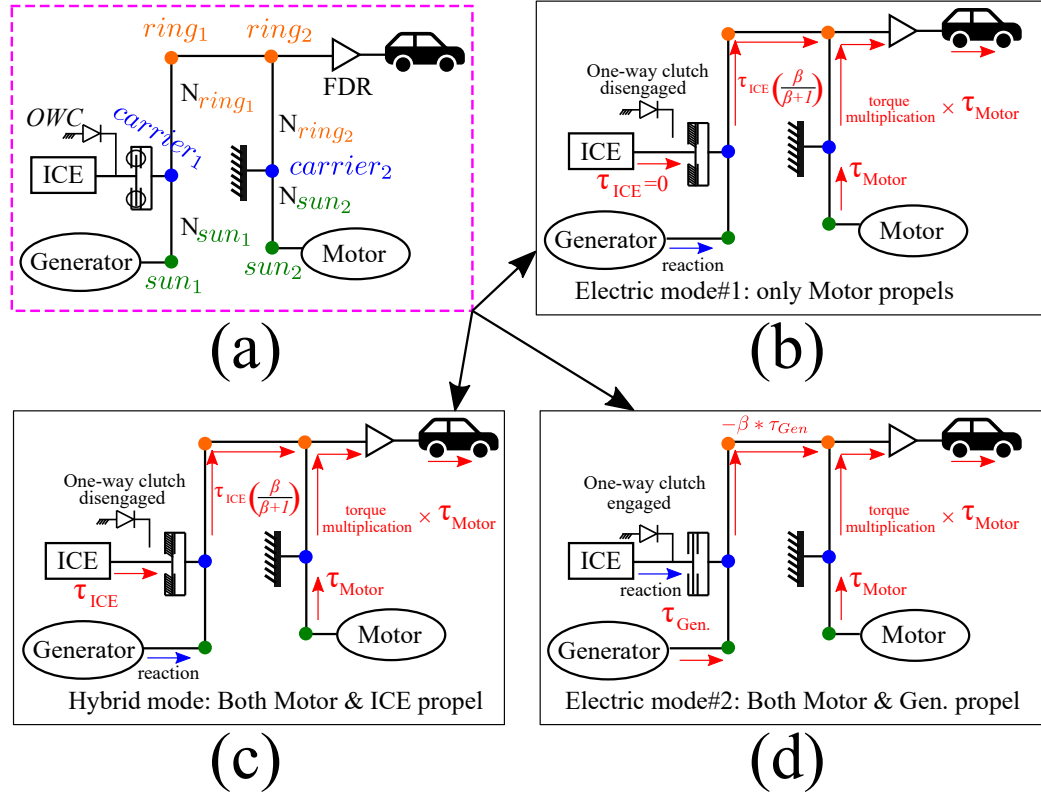


Figure 4.3: Schematic diagram of the three modes facilitated by the multi-mode powertrain chosen for this study

ICE model has been scheduled for future work. Tab. 4.1 tabulates the specifications of the powertrain components and vehicle components used in this study.

### Multi-mode e-VT modeling

The multi-mode e-VT with input-split PG set configuration [42] can facilitate three distinct electrified modes, i.e., single-motor EV mode, two-motor EV mode, and hybrid-electric mode with only a single PG set and an OWC between the planet carrier of PG set and crankshaft of the ICE. As depicted in Fig.4.3, the ICE is attached to the planet carrier of the PG set through a torque converter, generator is attached directly to the sun gear, and traction motor is attached to the ring gear

through a speed reduction gear and a counter-driven gear. This transaxle has almost same e-VT configuration as fourth generation THS and the presence of an OWC facilitates an additional dual-motor EV mode.

The kinematic relations among all the components of the e-VT transaxel is key for modeling the powertrain dynamics for each of the three distinct modes. The single-motor EV mode and HEV mode share the same powertrain dynamics and it is expressed through a set of kinematic relations as given by Eq.4.3.1a. The powertrain dynamics exhibited by the input-locked two-motor EV mode is modeled with a different set of kinematic relations as given by Eq.4.3.1b. The systematic method, as prescribed in [39], will be employed here to deduce the kinematic relations among the components of a multi-mode e-VT transaxle.

$$\{\tau\}_{HEV/EV1} = [A]_{HEV/EV1} \times \{\dot{\Omega}\}_{HEV/EV1} \quad (4.3.1a)$$

$$\{\tau\}_{EV2} = [A]_{EV2} \times \{\dot{\Omega}\}_{EV2} \quad (4.3.1b)$$

, where the generalized equation of obtaining state-dynamic ( $[A]$ ) matrix, torque ( $\{\tau\}$ ), and angular acceleration ( $\{\dot{\Omega}\}$ ) vectors corresponding to each mode is given by eq.(4.3.2) as follows:

$$\{\tau\}_{mode} = [M]_{mode} \{\tau_0\}; \{\dot{\Omega}\}_{mode} = [P]_{mode} \{\dot{\Omega}_0\} \quad (4.3.2)$$

$$[A]_{mode} = [M]_{mode} [A_0] [M]_{mode}^T$$

, where  $[A_0]$  is the  $4n \times 4n$  matrix and  $n$  is the number of planetary gear-set, generalized torque vector ( $\{\tau_0\}$ ), and generalized angular acceleration vector ( $\{\dot{\Omega}_0\}$ ) for

the e-VT configuration, used in this brief, are given by eq.(4.3.3) as follows:

$$A_0 = \begin{bmatrix} J_{out}+J_{r2} & 0 & 0 & 0 & 0 & 0 & 0 & -R_2 \\ 0 & J_{ice}+J_{c1} & 0 & 0 & 0 & 0 & R_1+S_1 & 0 \\ 0 & 0 & J_{gen}+J_{s1} & 0 & 0 & 0 & -S_1 & 0 \\ 0 & 0 & 0 & J_{mot}+J_{s2} & 0 & 0 & 0 & -S_2 \\ 0 & 0 & 0 & 0 & J_{r1} & 0 & -R_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & J_{c2} & 0 & 0 \\ 0 & R_1+S_1 & -S_1 & 0 & -R_1 & 0 & 0 & 0 \\ -R_2 & 0 & 0 & -S_2 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.3.3)$$

$$\{\tau_0\} = [\tau_{out} \ \tau_{ice} \ \tau_{gen} \ \tau_{mot} \ 0 \ 0 \ 0 \ 0]^T$$

$$\{\dot{\Omega}_0\} = [\dot{\Omega}_{out} \ \dot{\Omega}_{ice} \ \dot{\Omega}_{gen} \ \dot{\Omega}_{mot} \ \dot{\Omega}_{r1} \ \dot{\Omega}_{c2} \ F_1 \ F_2]^T$$

Here, it is required to elaborate the methodology used for deriving the powertrain dynamics of every mode facilitated by a multi-mode ePT with two PG-sets.

The rule for numbering the nodes of a two PG-set system are enumerated as follows:

- Always mark 1st to the node connected to the vehicle output.
- If ICE is there as a component, mark 2nd to the node connected to ICE.
- After marking these two nodes, mark 3rd to the first PG-set's node connected to any other prime-mover (traction motor/ generator).
- If both electric machines are on the same PG-set, then ring gear will priority in numbering.
- After all nodes with prime-movers are numbered, the remaining nodes are numbered as per priority. First PG-set gets priority over second, and ring gear, carrier, and sun gear are prioritized in descending order.

The powertrain configuration of Pacifica is exemplified in Fig.4.4 to show the numbering technique before proceeding for deriving the system dynamics of every powertrain

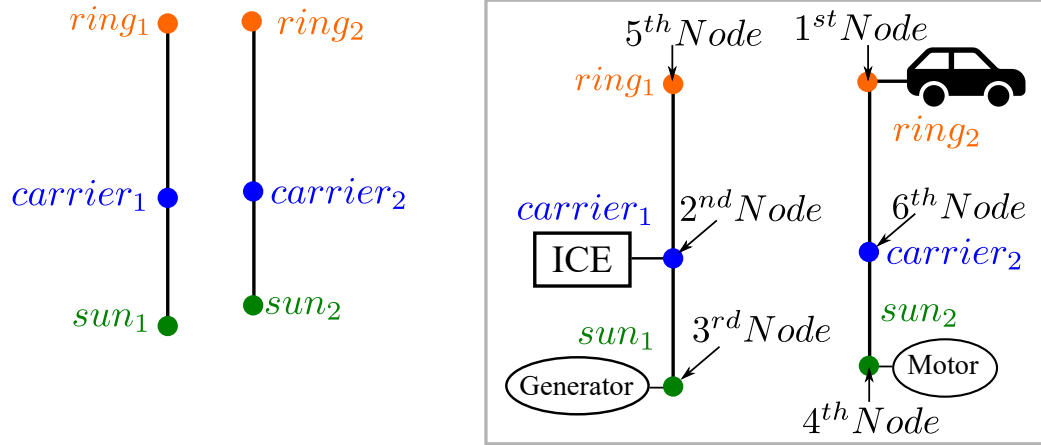


Figure 4.4: Numbering of nodes of a two-planetary gear set system with an example mode. The powertrain dynamics without any connection between nodes can be expressed by Eq.4.3.4.

$$A_0 \dot{\Omega}_0 = \begin{bmatrix} J & D \\ D^T & 0 \end{bmatrix} \begin{bmatrix} \dot{\Omega} \\ F \end{bmatrix} = \begin{bmatrix} \tau \\ 0 \end{bmatrix} = \tau_0 \quad (4.3.4)$$

$J$  is the diagonal matrix with a dimension of  $3n \times 3n$ , as shown in Eq.4.3.5 representing the inertia of the system. The diagonal elements of first six rows reflect the inertia terms of six nodes of the two planetary gear-set system. Among these six elements, first four represent the coupled inertia of the node and the prime-mover attached to it. It is evident from Fig.4.4 and Eq.4.3.3 that the remaining two nodes are not



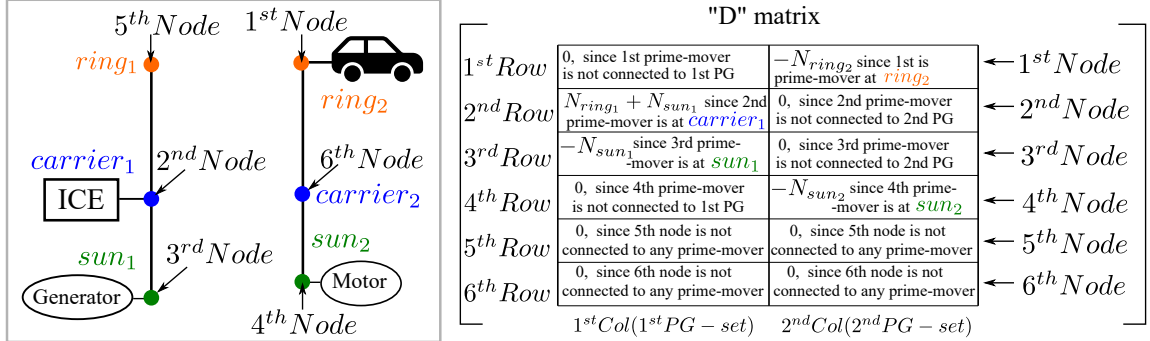


Figure 4.5: Methodology for obtaining constraints matrix, with an example.

connected to any prime-movers directly.

$$J = \begin{bmatrix} J_{out} + J_{1^{st}node} & 0 & 0 & 0 & 0 & 0 \\ 0 & J_{ice} + J_{2^{nd}node} & 0 & 0 & 0 & 0 \\ 0 & 0 & J_{gen} + J_{3^{rd}node} & 0 & 0 & 0 \\ 0 & 0 & 0 & J_{mot} + J_{4^{th}node} & 0 & 0 \\ 0 & 0 & 0 & 0 & J_{5^{th}node} & 0 \\ 0 & 0 & 0 & 0 & 0 & J_{6^{th}node} \end{bmatrix} \quad (4.3.5)$$

Next is how to determine the  $D$  and  $D^T$  matrices which are the upper-right  $3n \times n$  constraint matrix and bottom-left constraint matrix, respectively. The entries to the  $D$  matrix is governed by the connections between prime-movers and nodes of the PG-set. A clear guideline for constructing the  $D$  matrix is prescribed in [38]. Fig.4.5 elucidates the methodology of obtaining the  $D$  matrix. The first and second columns of the  $D$  matrix correspond to the first and second PG-sets, respectively. If any one of the four components is connected to the ring gear of a PG-set,  $D_{m,n}$  will be  $-N_{ring_m}$ , where  $m$  and  $n$  corresponds to index of PG-set and index of the component, respectively. Similarly, if any one of the four components is connected to the sun

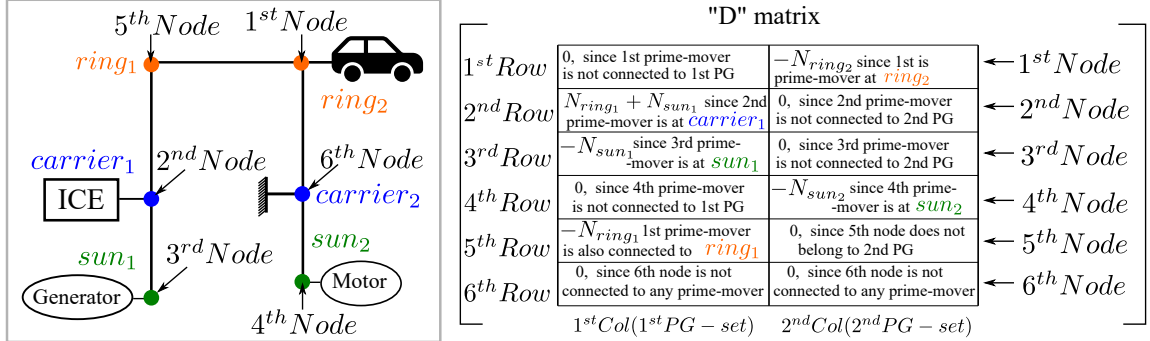


Figure 4.6: Constraints matrix for the powertrain architecture of Chrysler Pacifica with explanation.

gear of a PG-set,  $D_{m,n}$  will be  $-N_{sun_m}$ . Whereas, if any one of the four components is connected to the carrier of a PG-set,  $D_{m,n}$  will be  $(N_{sun_m} + N_{ring_m})$ .  $N_{sun}$ ,  $N_{ring}$ ,  $N_{carrier}$  are the teeth number of sun gear, ring gear, and carrier, respectively. In Eq.4.3.3,  $R_1$ ,  $R_2$ ,  $S_1$ , and  $S_2$  refer to  $N_{ring_1}$ ,  $N_{ring_2}$ ,  $N_{sun_1}$ , and  $N_{sun_2}$ , respectively.

The  $D$  matrix for the powertrain architecture of Chrysler Pacifica is elaborated in Fig.4.6. There is another piece of the puzzle, i.e., transition matrices ( $M$  and  $P$  matrices), for deriving the system dynamics corresponding to every distinct mode of the powertrain architecture. The derivation of  $M$  and  $P$  matrices are equally important to obtain the final form of system dynamics for every mode. Every distinct mode has a set of distinct  $M$  and  $P$  matrices. Transition matrices  $M$  and  $P$  are derived as per the clutch engagement.

$M$  is initialized as a  $4n \times 4n$  identity matrix, the same dimension as  $A_0$ . When the  $i$ th PG-set node is connected with the  $j$ th PG-set node, assuming  $i < j$ , the processes shown in Eqs.4.3.6, 4.3.7, and 4.3.8 are applied to modify the  $M$  matrix. If the  $i$ th node is clutched to ground,  $i$ th row will be eliminated. Fig.4.7 depicts the process to

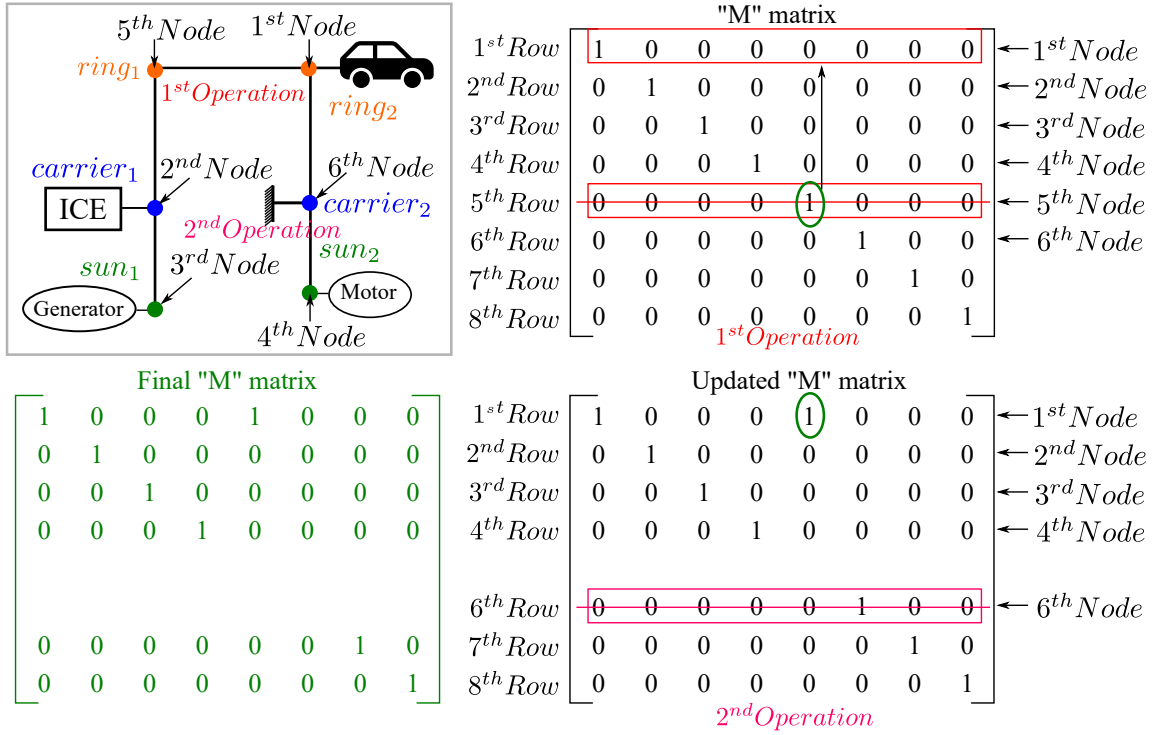


Figure 4.7: Transition "M" matrix for the powertrain architecture of Chrysler Pacifica for the HEV and single-motor all-electric modes.

obtain  $M$  matrix for both single-motor EV mode and hybrid-electric mode.

$$i^{th}row = i^{th}row + j^{th}row \quad (4.3.6)$$

$$j^{th}row = [] \quad (4.3.7)$$

$$i^{th}row = [] \quad (if \ i^{th} \ node \ is \ clutched \ to \ ground) \quad (4.3.8)$$

$P$  is also initialized as a  $4n \times 4n$  identity matrix, the same dimension as  $A_0$ . When

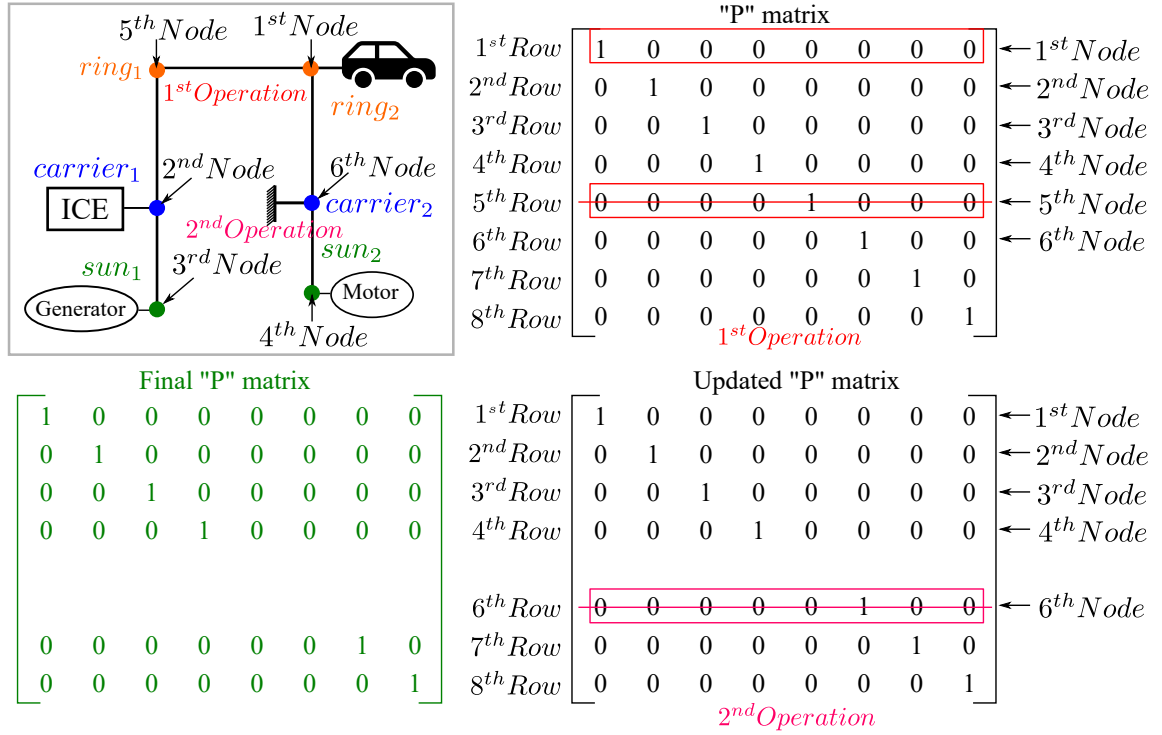


Figure 4.8: Transition “P” matrix for the powertrain architecture of Chrysler Pacifica for the HEV and single-motor all-electric modes.

the  $i$ th PG-set node is connected with the  $j$ th PG-set node, assuming  $i < j$ , the processes shown in Eqs.4.3.7 and 4.3.8 are applied to modify the  $P$  matrix. If the  $i$ th node is clutched to ground,  $i$ th row will be eliminated. Fig.4.8 depicts the process to obtain  $P$  matrix for both single-motor EV mode and hybrid-electric mode. Finally,

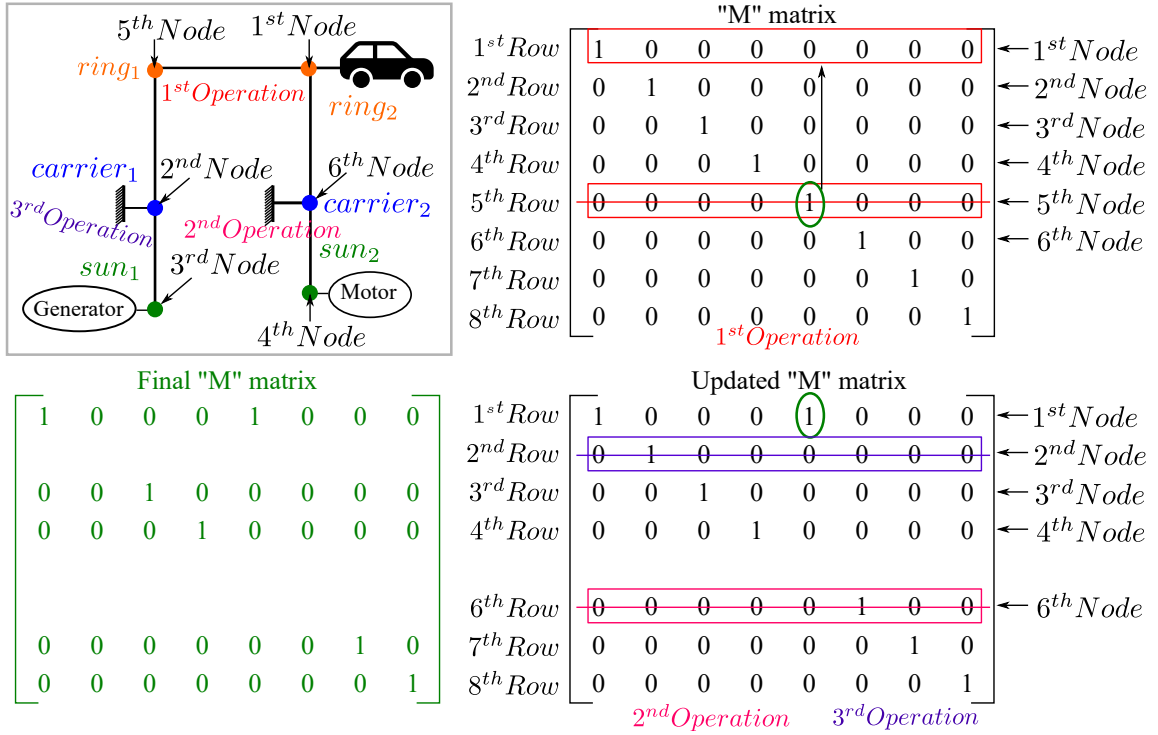


Figure 4.9: Transition "M" matrix for the powertrain architecture of Chrysler Pacifica for the two-motor all-electric mode.

the  $M_{HEV/EV1}$  and  $P_{HEV/EV1}$  matrices can be expressed as follows:

$$M_{HEV/EV1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad P_{HEV/EV1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.3.9)$$

Similarly, the  $M$  and  $P$  matrices for two-motor all-electric mode is presented in Fig.4.9 and 4.10, respectively. The final expressions for them are provided in the following

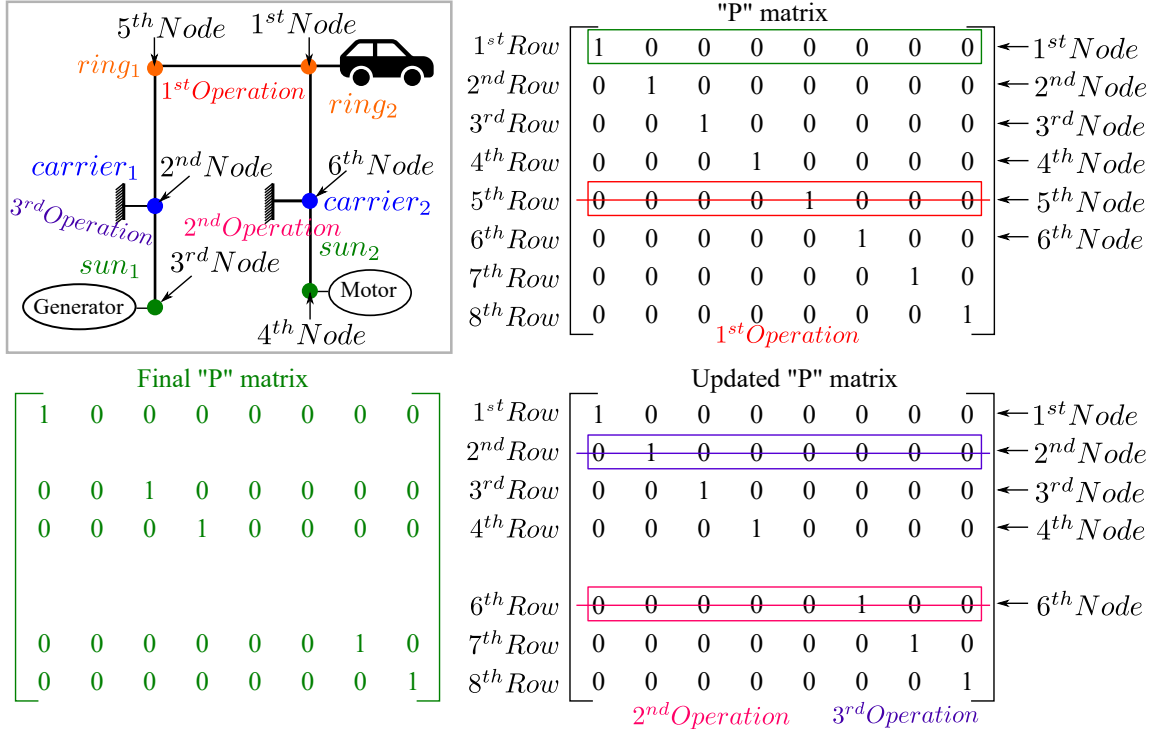


Figure 4.10: Transition "P" matrix for the powertrain architecture of Chrysler Pacifica for the two-motor all-electric mode.

equation.

$$M_{EV\#2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad P_{EV\#2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.3.10)$$

Finally, the required expanded kinematic relations for both the distinct modes can be obtained through substituting eq.(4.3.3), (4.3.9), and (4.3.10) in eq.(4.3.2) and they

are given as follows:

$$\begin{bmatrix} \dot{\Omega}_{out} \\ \dot{\Omega}_{ice} \\ \dot{\Omega}_{gen} \\ \dot{\Omega}_{mot} \\ F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} J_{outeq} & 0 & 0 & 0 & -R_1 & -R_2 \\ 0 & J_{iceeq} & 0 & 0 & R_1 + S_1 & 0 \\ 0 & 0 & J_{geneq} & 0 & -S_1 & 0 \\ 0 & 0 & 0 & J_{moteq} & 0 & -S_2 \\ -R_1 & R_1 + S_1 & -S_1 & 0 & 0 & 0 \\ -R_2 & 0 & 0 & -S_2 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \tau_{out} \\ \tau_{ice} \\ \tau_{gen} \\ \tau_{mot} \\ 0 \\ 0 \end{bmatrix} \quad (4.3.11)$$

$$\begin{bmatrix} \dot{\Omega}_{out} \\ \dot{\Omega}_{gen} \\ \dot{\Omega}_{mot} \\ F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} J_{outeq} & 0 & 0 & -R_1 & -R_2 \\ 0 & J_{geneq} & 0 & -S_1 & 0 \\ 0 & 0 & J_{moteq} & 0 & -S_2 \\ -R_1 & -S_1 & 0 & 0 & 0 \\ -R_2 & 0 & -S_2 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \tau_{out} \\ \tau_{gen} \\ \tau_{mot} \\ 0 \\ 0 \end{bmatrix} \quad (4.3.12)$$

While eq.(4.3.11) defines the kinematic relations for single-motor EV mode and HEV mode, eq.(4.3.12) can be applied to the dual-motor EV mode.  $\dot{\Omega}_{gen}$ ,  $\dot{\Omega}_{ICE}$ ,  $\dot{\Omega}_{mot}$ , and  $\dot{\Omega}_{out}$  are the angular accelerations of generator, ICE, traction motor, and ring gear of the PG set respectively.  $J_{outeq} = (J_{out} + J_{ring_1} + J_{ring_2})$  is the cumulative MMI of vehicle, ring gear#1 and counter-driven gear at the transmission output.  $J_{iceeq} = (J_{ice} + J_{carrier_1})$ ,  $J_{geneq} = (J_{gen} + J_{sun_1})$ , and  $J_{moteq} = (J_{mot} + J_{sun_2})$ .  $J_{gen}$ ,  $J_{ice}$ ,  $J_{mot}$ ,  $J_{ring_2}$ ,  $J_{ring_1}$ ,  $J_{carrier_1}$ ,  $J_{sun_1}$ , and  $J_{sun_2}$  are the MMI of generator, ICE, traction motor, counter-driven gear, ring gear#1, carrier#1, sun gear#1, and the pinion attached to motor shaft respectively.  $\tau_{gen}$ ,  $\tau_{ICE}$ ,  $\tau_{mot}$ ,  $\tau_{out}$  are the torques applied by generator, ICE, traction motor, and transmission output respectively.  $\beta$

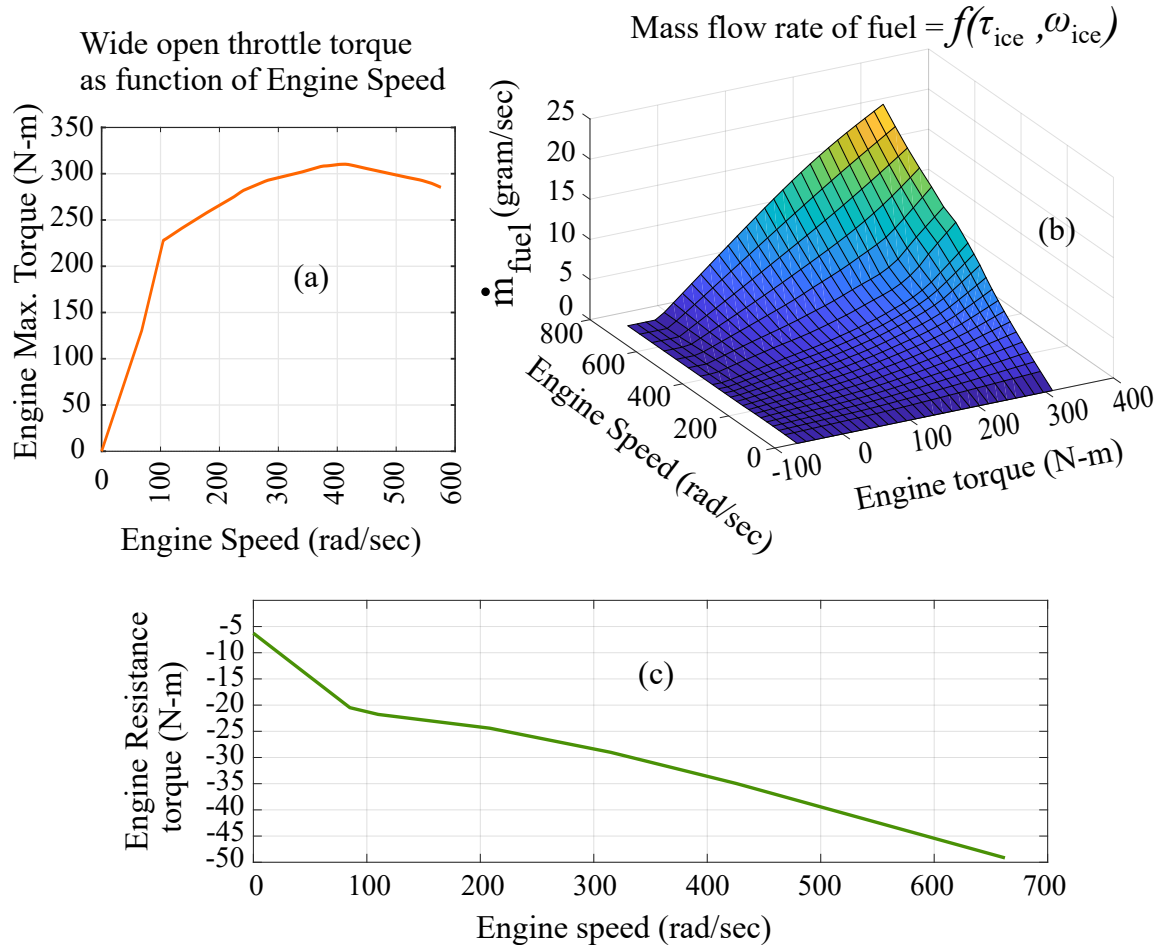


Figure 4.11: Mass flow rate of fuel consumption, WOT torque, and friction torque.

is the ratio of number of teeth of ring gear to sun gear ( $\beta = \frac{R}{S}$ ).

### Modeling of Internal Combustion Engine

A low fidelity 3.3 Lt spark-ignition (SI) ICE is modeled through three most essential engine characteristic look-up tables, i.e., WOT torque, mass flow rate (MFR), and friction torque. The MFR of fuel from the fuel injectors of an SI engine is calculated



in terms of fuel-air equivalent ratio (FAER) as follows:

$$\begin{aligned}
 \dot{m}_{fuel} &= FAER * \frac{\dot{m}_{air}(\omega_{ICE_{actual}}, \tau_{ICE_{request}})}{\text{Stoichiometric air fuel ratio}} \\
 &= FAER * \frac{\dot{m}_{air}(\omega_{ICE_{actual}}, \tau_{ICE_{request}})}{15.0461} \\
 &= 1 \times \frac{\dot{m}_{air}(\omega_{ICE_{actual}}, \tau_{ICE_{request}})}{15.0461}
 \end{aligned} \tag{4.3.13}$$

, where FAER is the fuel-air equivalent ratio which depends on the control algorithm of engine control unit (ECU). The value of FAER can range from 0.02 to 10. The reference value of FAER is 1 (one) which corresponds to stoichiometric fuel-air mixture. FAER value higher than 1 and lower than 1 refers to rich and lean fuel-air mixture respectively. Generally MFR of air is a direct function actual engine speed and torque request and ECU always tries to achieve the reference value of FAER which is 1.

The high-frequency engine ripple torque is replaced by low-frequency mean-value engine resistance torque, as shown in Fig.4.11.c. Expression of high-frequency engine ripple torque, a function of instantaneous engine speed and indicated torque, is given in Eq.4.3.14 [140].

$$\tau_{fr_{ice}} = [C_0 + C_1(\dot{\theta}_{ice}) + C_2(\dot{\theta}_{ice})^2] \times \left( \frac{V_d}{4\pi} + \frac{\tau_{ind}(\theta_{ice})}{40} \right) \tag{4.3.14}$$

, where  $\tau_{fr_{ice}}$ ,  $C_{0,1,2}$ ,  $\theta_{ice}$ ,  $V_d$ ,  $\tau_{ind}(\theta_{ice})$  are friction torque of engine, approximating factors, engine crank angle, engine volume, and indicated torque.

## EMG modeling

Low-fidelity models of both generator and traction motor are developed with a simple look-up table approach. As shown in the Fig.4.12 and Fig.4.13, the maximum torque

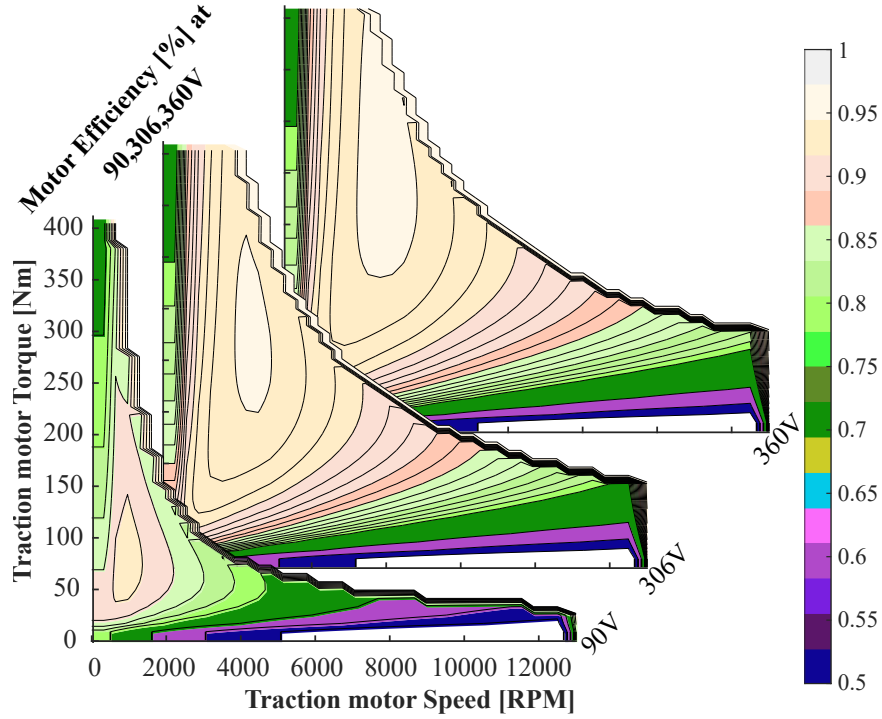


Figure 4.12: Efficiency contour plots of traction motor at different voltages.

( $\tau_{MG_{max.}}$ ) curves and efficiency ( $\eta_{MG}$ ) table can be represented as following relations:

$$\begin{aligned}\tau_{MG_{max.}} &= f(\omega_{MG}, Voltage_{MG}) \\ \eta_{MG} &= f(\omega_{MG}, \tau_{MG}, Voltage_{MG})\end{aligned}\tag{4.3.15}$$

These curves are obtained through the motor curve creator tool in Simcenter AMESim®.

### High-Voltage Battery modeling

A low fidelity HVB model is adequate for developing an EMS control strategy for HEPT. Highly accurate SOC estimation through complex nonlinear battery modeling technique [141] might be useful during validation stage prior to HIL simulation. At

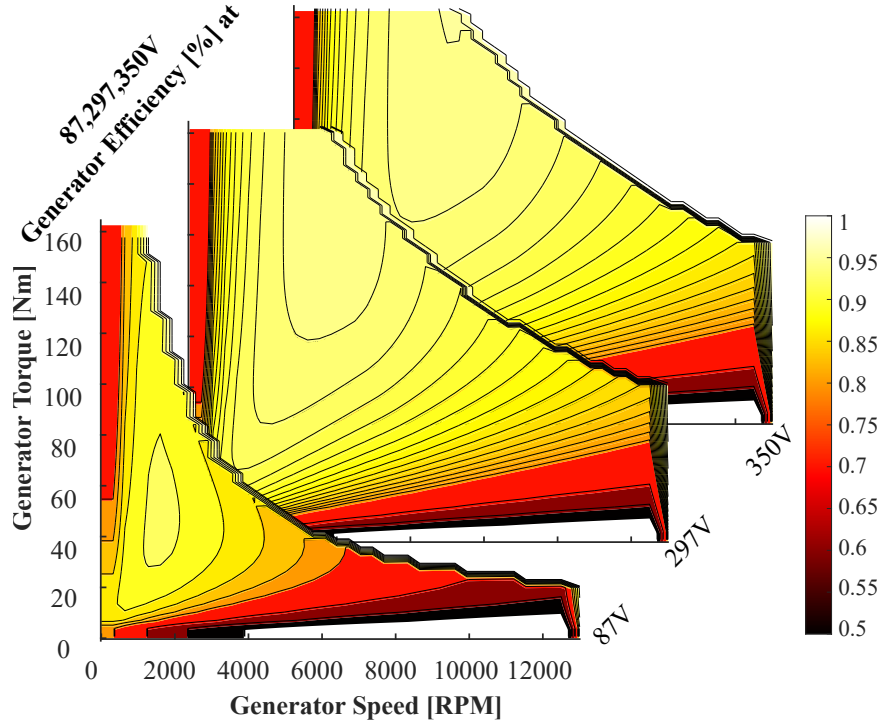


Figure 4.13: Efficiency contour plots of generator at different voltages.

this initial control development stage, the HVB is simply modeled with a open circuit voltage (OCV) in series with internal resistance (IR) of the HVB. SOC dynamics is given by the following equation:

$$\frac{d}{dt}SOC(t) = -\frac{OCV - \sqrt{OCV^2 - 4P_{batt} * IR}}{2IR * Q_{batt}} \quad (4.3.16)$$

, where  $P_{batt}$  and  $Q_{batt}$  are power and overall capacity of the HVB respectively.

#### 4.3.4 ECMS Controller

The ECMS controller is the most crucial block in this entire simulation platform since it decides the operating mode of the powertrain, and if the mode has controllability,

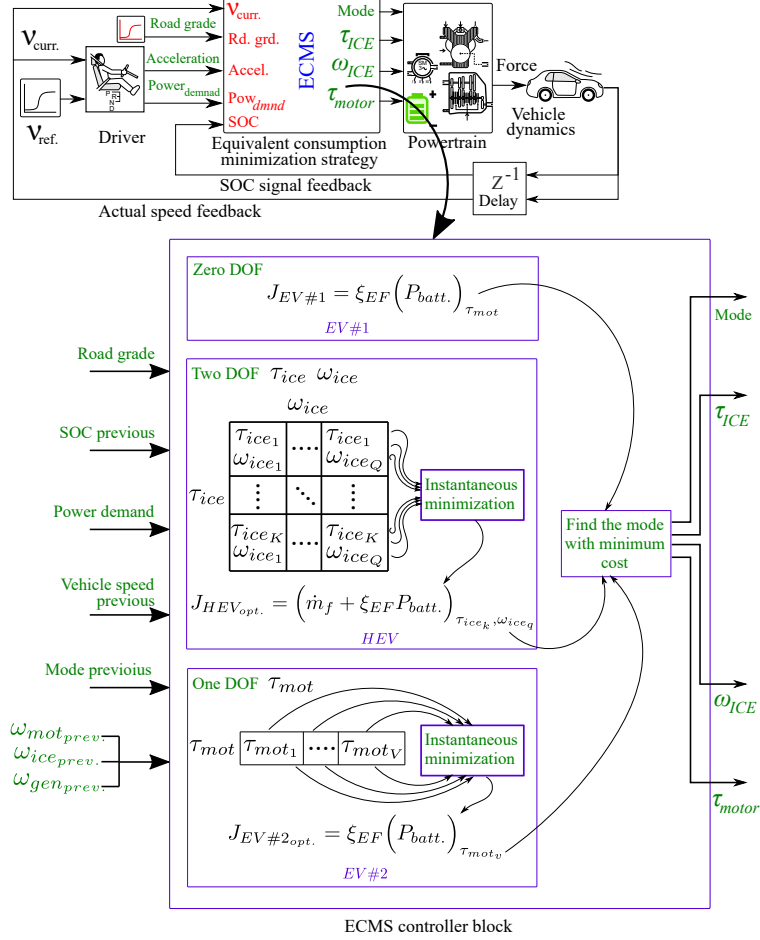


Figure 4.14: Overview of the ECMS controller block in a model-in-the-loop simulation with a multi-mode electrified powertrain.

the ECMS controller will decide operating points of the controllable prime-movers through instantaneous optimization. The overview of the ECMS controller is given in Fig.4.14. Assisted by the Fig.4.14, it will be easier to comprehend the algorithm of an ECMS prescribed for a multi-mode ePT. The algorithm is as follows:

---

**Algorithm 1:** Equivalent Consumption Minimization Strategy for Multi-mode Electrified Powertrain

---

**Input :**  $\tau_{rqst_{drvr}}, SOC_{prev}, \omega_{veh_{prev}}, \omega_{ice_{prev}}, \omega_{mot_{prev}}, \omega_{gen_{prev}}$  &  $\omega_{rqst_{drvr}}$   
**Output :**  $U(t)$ , where  $U = \{mode_{optim.}, \Psi_{mode_{optim.}}\}$ , where  $mode = \{1, 2, 3\}$ ,  
 $\Psi_1 = \{\}$ ,  $\Psi_2 = \{\tau_{ice_{cmnd}}, \omega_{gen_{cmnd}}\}^T$ ,  $\Psi_3 = \{\tau_{mot_{cmnd}}\}$ ,  
where  $1 \equiv EV\#1$ ,  $2 \equiv HEV$ ,  $3 \equiv EV\#2$   
**Initialization :**  $\tau_{out} = \tau_{rqst_{drvr}}$ ,  
 $\omega_{out} = \omega_{rqst_{drvr}}, SOC = SOC_{prev}; Volt = Volt_{prev}$   
**for**  $mode = i, i \in \{1, 2, 3\}$  **do**  
    sweep through feasible combinations of control variables in mode#i ( $\Psi_i$ )  
    **for** each feasible combinations of  $\Psi_i$  **do**  
        **compute:**  $\omega_{ice}, \tau_{ice}, \omega_{gen}, \tau_{gen}, \omega_{mot}$ , and  $\tau_{mot}$  corresponding to  $\Psi_i$   
        (use kinematic relationship for the given mode, eq.(2.3.4) for EV#1  
        & HEV and eq.(2.3.5) for EV#2)  
        **compute:**  $\dot{m}_{fuel}$  corresponding to  $\omega_{ice}, \tau_{ice}$   
        **compute:**  $Pow_{mot} = (Pow_{mot_{useful}} + Loss_{mot})$   
        **compute:**  $Pow_{gen} = (Pow_{gen_{useful}} + Loss_{gen})$   
        **compute:**  $Pow_{batt} = (Pow_{mot} + Pow_{gen} + Loss_{mot})$   
        **compute:**  $J_{\Psi_i} = (\dot{m}_{fuel} \times LHV) + \lambda \times Pow_{batt}$   
     $\Psi_{i,optim.} = \arg \min_{\Psi_i} \|(J_{\Psi_i})\|$   
 $mode_{optim.} = \min(J_{mode\#1}, J_{mode\#2}, J_{mode\#3})$   
 $U(t) = \{mode_{optim.}, \Psi_{mode_{optim.}}\}$

---

The discretizations used for  $\tau_{ice}, \omega_{ice}$ , and  $\tau_{mot}$  are as follows:

$$\tau_{ice} = [13.5 : 15 : 310.5] (N - m) \quad (4.3.17a)$$

$$\omega_{ice} = [52 : 20 : 575] (rad/sec) \quad (4.3.17b)$$

$$\tau_{mot} = \left[ [0 : (f(\omega_{mot}, Voltage_{mot}))/15 : f(\omega_{mot}, Voltage_{mot})] (N - m) \right] \quad (4.3.17c)$$

The cost function is a key characteristic of an ECMS. In this dissertation, only fuel consumption minimization is taken into consideration. To encompass the effect of both fuel consumption and emission, the ECMS controller will no longer be able to

minimize the fuel consumption because of the fact that the lowest fuel consumption region of the ICE map does not necessarily yield lowest emissions. Another set of look-up tables or empirical models for emissions are indispensable to optimize the cost function comprised of instantaneous equivalent fuel consumption and tailpipe emissions. However, another feature can be easily augmented to the existing ECMS with equivalent consumption minimization. Drivability and mode-shift's feasibility might not seem the primary objective of the basic ECMS at this model-in-the-loop simulation stage. However, in later stage of the controller development, inevitably such a secondary objective becomes very practical requirement to be fulfilled by the ECMS.

Drivability is a subjective matter, and is defined as the driver's perception about driving experience. There are various metrics available for measuring the drivability score of an ePT. However, a simple metric, i.e., number of ICE activation has been used in this work to measure and improve the drivability. With the following cost function, the ECMS controller tend to switch pretty frequently between ICE activation and deactivation.

$$J_{\Psi_i} = (\dot{m}_{fuel} \times LHV) + \lambda \times Pow_{batt} \quad (4.3.18)$$

Frequent activation of ICE implies to frequent deactivation, which brings not only just drivability issue but more fuel consumption during ICE cranking. Once cranked, the ICE is expected to run at least for ten seconds at a stretch to reach steady-state operation for better efficiency. Although the transients of ICE has not been modeled accurately in this work (scheduled for future work), it is ensured that ICE is not deactivated until at least ten seconds of operation to emulate the practical feasibility constraint. To prevent ICE deactivation immediately after its activation, an extra

penalty cost has been added to the all-electric modes consistently for the next ten seconds, as shown below in Eq.4.3.19 and 4.3.20

$$J_{\Psi_i}|_{EV\#1} = \left\{ \begin{array}{ll} \lambda \times Pow_{batt}, & \text{for } Mode_{prev} == EV\#1||EV\#2 \\ \lambda \times Pow_{batt} + Penalty_{ice}, & \text{for } Mode_{prev} == HEV \&\& counter \leq 10sec \end{array} \right\} \quad (4.3.19)$$

$$J_{\Psi_i}|_{EV\#2} = \left\{ \begin{array}{ll} \lambda \times Pow_{batt}, & \text{for } Mode_{prev} == EV\#1||EV\#2 \\ \lambda \times Pow_{batt} + Penalty_{ice}, & \text{for } Mode_{prev} == HEV \&\& counter \leq 10sec \end{array} \right\} \quad (4.3.20)$$

, where  $Penalty_{ice}$  is the ICE-deactivation penalty within the period of ten seconds after ICE-activation. The variable  $counter$  tracks the difference between the current time-step and time-step when ICE was last activated. Similarly, ICE-activation should be prohibited immediately after deactivation of ICE. Hence, a similar penalty should be added to the cost of hybrid-electric mode when the vehicle is running in any one of the all-electric modes. The modified cost function of hybrid-electric mode will prohibit the ECMS from activating the hybrid-electric mode during five seconds of its post-deactivation phase.

$$J_{\Psi_i}|_{HEV} = \left\{ \begin{array}{ll} \begin{array}{l} (\dot{m}_{fuel} \times LHV) \\ + \lambda \times Pow_{batt} \end{array}, & \text{for } Mode_{prev} == HEV \\ \begin{array}{l} (\dot{m}_{fuel} \times LHV) \\ + \lambda \times Pow_{batt} + Penalty_{EV} \end{array}, & \text{for } Mode_{prev} == EVs \&\& counter \leq 5sec \end{array} \right\} \quad (4.3.21)$$

Feasibility of mode-shift events is a crucial factor for any EMS strategy dealing with a multi-mode ePT. While the mode-shift events and their feasibility are all together a separate domain and will be dealt judiciously in the next chapter, a glimpse of the implementation of feasibility criteria is shown here. Inspired by the work presented in [142], the original equivalent cost structure, given in Eq.4.3.18, has been modified as below to yield feasible mode-shift commands from ECMS controller.

$$(J)_{equi}^{modif.} = \dot{m}_{fuel} * (LHV) + \lambda * Pow_{batt.} + \nu * P_{shift} \quad (4.3.22)$$

, where

$$\begin{aligned} P_{shift_{mode}} &= \mu_1 * \frac{1}{2} J_{ice} [(\omega_{ice}^{t+1})^2 - (\omega_{ice}^t)^2] \\ &+ \mu_2 * \frac{1}{2} J_{gen} [(\omega_{gen}^{t+1})^2 - (\omega_{gen}^t)^2] \\ &+ \mu_2 * \frac{1}{2} J_{mot} [(\omega_{mot}^{t+1})^2 - (\omega_{mot}^t)^2] \end{aligned}$$

$\nu$  is a tuning factor, and a value of 0.03 is selected for this work. Different values have been assigned to tuning factors  $\mu_1$ , and  $\mu_2$  corresponding to different types of mode-shift and the assigned values are tabulated in Tab.4.2. The feasibility constraint of mode-shift event does not let the ECMS controller choose operating points, in two consecutive time-steps for the ePT, having drastically different numerical value. For example, if the ICE is operating at 250 *rad/sec* at current time-step, the feasibility constraint will not let the ECMS controller choose 350 *rad/sec* in the next time-step. It is noteworthy to mention that the time-step used for the simulation in this work is 0.1 *seconds*.



## 4.4 Summary

This chapter elucidated the implementation of ECMS as the control strategy for an EMS of multi-mode ePT. Implementation of ECMS requires the derivation of the powertrain dynamics in every operating mode as a first priority. Therefore, the derivation has been illustrated in detail. It is noteworthy to mention that brute force search method has been adopted for implementing the instantaneous optimization in this dissertation. Brute force optimization is feasible in real-time as long as the discretization is coarse or boundaries of control variables are small. Brute force search is the safest solution if the cost function is non-convex in nature. But, the computational time for brute force exceeds real-time threshold as discretization becomes fine with large boundaries of the control variables. Hence, non-convex optimization techniques, for 2-D search-space, such as stochastic gradient descent or gradient projection will be deployed in future to increase the accuracy of optimization and to meet the real-time computation constraint.

Table 4.1: Specification of the vehicle and cardinal components

Component	Parameter	Value
Vehicle	Mass (m)	2510 kg
	Rolling resistance co-efficient	$f_1 = 0.006, f_2 = 0.0001$
	Vehicle inertia ( $J_{out}$ )	298.08 kg-m <sup>2</sup>
	Wheel radius ( $r_{wheel}$ )	0.3548 m
	Vehicle height	1.75 m
	Vehicle width	1.48 m
	Engine	Inertia ( $J_{ice}$ )
Compression ratio		
Max. speed ( $\omega_{ice_{max}}$ )		5500 RPM
Max. torque ( $\tau_{ice_{max}}$ )		310 Nm at 4100 rpm
Volume displacement		4 Cyl., 2.0152 liters
Max. power		175 kWatt at 5000 rpm
Traction motor	Inertia ( $J_{mot}$ )	0.0028 kg-m <sup>2</sup> ,
	Max. speed, Max. Volt.	12500 RPM, 414 V
	Max. torque ( $\tau_{mot_{max}}$ )	410 Nm at [0:2240]
	Max. power	96 kWatt at 2240 rpm
Generator	Inertia ( $J_{gen}$ )	0.0024 kg-m <sup>2</sup>
	Max. speed, Max. Volt.	12500 RPM, 350 V
	Max. torque ( $\tau_{gen_{max}}$ )	162.5 at [0:4167] rpm
	Max. power	70 kWatt at 4167 rpm
Battery	Max. capacity	43 Ah
	Max. discharge power	60 kWatt
	Max. recharge power	40 kWatt
	No. of cells in series	100
	Max. voltage	414 V

Table 4.2: Tuning factor for mode-shift penalty term

Current mode			
Previous mode		Fully electric	Hybrid electric
	Fully electric		$\mu_1 = 0, \mu_2 = 1$
Hybrid electric		$\mu_1 = 0.1, \mu_2 = 0.5$	$\mu_1 = 1, \mu_2 = 1$

## Chapter 5

**In Pursuit of a Novel Strategy:  
Overall System Loss Minimization  
Strategy instead of Engine's  
Efficiency Maximization**

## 5.1 Introduction

As concluded in the previous chapter, researches are going on for finding new derivatives of optimal ECMS to dissolve the sufficient condition related to a unique value of EF yielding global optimal performance. In this chapter, a new instantaneous optimization-based strategy, which is similar to ECMS, is proposed to minimize the overall powertrain loss unlike ECMS which focuses on just maximizing ICE's efficiency. The LMS is articulated with an ambition of outperforming optimal ECMS, given that LMS does know the drive cycle a priori. However, LMS also tend to switch among different powertrain modes quite frequently alike ECMS. The chapter will substantiate the effectiveness of LMS in achieving almost same performance as obtained by optimal ECMS, but not enough dexterity to outperform optimal ECMS in model-in-the-loop (MIL) simulation.

## 5.2 Fundamentals of Loss Minimization Strategy

The LMS is based on the principle of minimizing the loss of the entire powertrain rather than just considering the effects of only the ICE while minimizing fuel consumption. Minimizing the cumulative energy loss of all the subsystems is equivalent to minimizing the fuel consumption of the vehicle [143]. Therefore, utilizing the concept from [144], the optimal control problem statement can be defined as:

$$[u_{ice}^{opt}(t), u_{mot/gen}^{opt}(t)] = \arg \min_{u_{ice}(t), u_{mot/gen}(t)} [C_t], \quad t = 1, \dots, T_f \quad (5.2.1)$$

, where

$$\begin{aligned}
 C_t &= \sum_{m \in \mathcal{M}} E_{in,m} - E_{out,m} + \lambda E_{hvb} \\
 &= \sum_{m \in \mathcal{M}} \left( u_m(t) - y_m(t) \right) \Delta t + \lambda u_{hvb}(t) \Delta t
 \end{aligned} \tag{5.2.2}$$

$$\begin{aligned}
 E_{req}(t) &= E_{ice}(t) + E_{gen}(t) + E_{mot} \\
 SOC_{min} &< SOC(t) < SOC_{max} \\
 \text{subject to} \quad 0 &\leq E_{ice}(t) \leq E_{ice,max}(t) \\
 E_{gen,min}(t) &\leq E_{gen}(t) \leq E_{gen,max}(t) \\
 E_{mot,min}(t) &\leq E_{mot}(t) \leq E_{mot,max}(t)
 \end{aligned} \tag{5.2.3}$$

, where  $E_{in}$  represents the amount of energy entering into a component,  $E_{out}$  represent energy flowing out and  $E_{hvb}$  represents the energy provided by the HVB at a given time-step.  $\lambda$  is the EF between chemical energy of fuel and electrical energy of HVB.  $\Delta t$  is the sampling time of online simulation.  $\mathcal{M}$  is the set containing all the mechanical and electrical components in the electrified powertrain. The system, as presented in Fig.5.1, describes the power flow in the current given configuration.  $u$  and  $y$  refers to input and output power flow for all the energy exchanging components.  $V_1$  is the output power demand of the electrified powertrain,  $V_2$  represents external energy exchange for the battery,  $X_{ice}$  represents fuel consumption of the ICE, and  $X_{hvb}$  represents SOC of the EESS which is a HVB for this application. The energy exchanging nodes are referred by  $n_k$  for  $k \in [1, 2, 3, 4]$ . Fig. 5.1 also exemplifies a clear guideline for deciphering the power-flow diagram for any given powertrain configuration. Node 1 ( $n_1$ ) represents the energy exchange between ICE, generator (EMG1), and transmission output. According to this particular configuration, output power

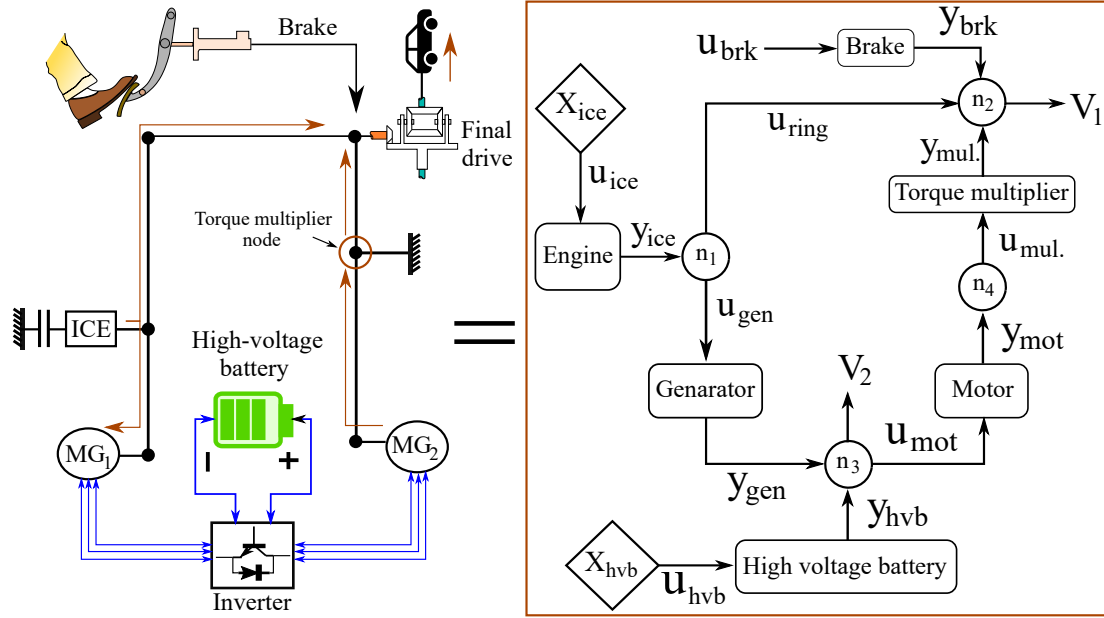


Figure 5.1: Power flow in multi-mode power-split electrified powertrain

of ICE ( $y_{ice}$ ) is divided into two paths, i.e., towards the final drive and towards the generator. Node 2 ( $n_2$ ) represents the junction where powers from ICE and traction motor (EMG2) coalesce, vehicles power demand flows outward, and equivalent brake power flows inward. Node 3 ( $n_3$ ) represents where powers of generator, motor, HVB, and external electrical power interact. EESS of an HEV does not get recharged through external electrical power source and hence  $V_2$  will be zero for this powertrain application. Furthermore, since there is only conceptual power aggregation at each node, the power should be balanced across every node in fig. 5.1 as per Eq.5.2.4. The power balance for nodes  $n_1$ ,  $n_2$ ,  $n_3$ , and  $n_4$  is given by Eq.5.2.4a, Eq. 5.2.4b, Eq.5.2.4c, and Eq.5.2.4d respectively.

$$0 = u_{ring} + u_{gen} - y_{ice} \quad (5.2.4a)$$

$$0 = V_1 - y_{mul.} - u_{ring} - y_{brk} \quad (5.2.4b)$$

$$0 = V_2 + u_{mot} - y_{gen} - y_{hvb} \quad (5.2.4c)$$

$$0 = u_{mul.} - y_{mot} \quad (5.2.4d)$$

, where  $u_{ring}$  represents the fraction of ICE power flowing towards transmission output through both ring gear#1 and counter-driven gear.  $u_{ice}$ ,  $u_{gen}$ ,  $u_{mot}$ ,  $u_{hvb}$ ,  $u_{brk}$ , and  $u_{mul.}$  represent input power flow to ICE, generator, motor, HVB, mechanical brake, and torque multiplier respectively. Similarly,  $y_{ice}$ ,  $y_{gen}$ ,  $y_{mot}$ ,  $y_{hvb}$ ,  $y_{brk}$ , and  $y_{mul.}$  represent output power flow from the aforementioned components in the same order. As per Eq.5.2, the instantaneous cost  $C_t$  of the electrified powertrain which needs to be minimized at every time-step ( $\Delta t$ ) can be articulated as:

$$C_t = (u_{ice} - y_{ice} + u_{gen} - y_{gen} + u_{mot} - y_{mot} + u_{mul.} - y_{mul.} + u_{hvb} - y_{hvb} + u_{brk} - y_{brk} + \lambda u_{hvb}) \Delta T \quad (5.2.5)$$

By substituting Eq.5.2.4a, Eq.5.2.4b, Eq.5.2.4c, and Eq.5.2.4d in Eq.5.2.5 and setting  $u_{br} = 0$ , since all the energy supplied for braking can be considered as loss, the instantaneous cost can be re-written as Eq.5.2.6.

$$C_t = (u_{ice} + (1 + \lambda)u_{hvb} - V_1) \Delta t \quad (5.2.6)$$

It can be noticed, that the Eq.5.2.6 represents the loss structure of the ECMS since  $V_1$  is the known output power demand at every time-step. As mentioned earlier, the biggest drawback of ECMS is that it requires tuning of  $\lambda$  for each unique drive-cycle and vehicle topology to yield near optimal results. Furthermore, in the case of an HEV, the ECMS guarantees charge sustaining results for online application only with a single optimal constant value of  $\lambda$  which should be pre-calculated with prior information about the entire drive cycle. Slight deviation from the optimal value will yield sub-optimal results and can result in violation of charge sustainability criteria. Since Eq.5.2 has the same cost structure as of the ECMS, it is implicit that LMS will also suffer from similar disadvantages.

### 5.3 Revised Cost Structure

One of the biggest issues with solving optimal problem pertaining to the EMS of an HEV is that it has inequality constraints over the state variable which needs to be satisfied. For an HEV, the HVB must operate in charge sustaining mode. It implies the final SOC of HVB must be within a pre-specified range at the end of the drive cycle. Generally, it is set to be within 1% around the initial SOC value. Since Eq.5.2 is highly sensitive to the value of  $\lambda$  it should be evaluated judiciously at each time-step. Literature is crowded with many methods specifying clear guideline for tuning the EF ( $\lambda$ ). This paper focuses on developing a new cost structure for the IOC of EMS pertaining to HEV while taking charge sustainability strictly into consideration. It is noteworthy to mention that the term  $E_{in,m} - E_{out,m}$  in Eq.5.2 basically represents the amount of energy being lost across each component. Henceforth, the lost energy will be represented by  $L_{\mathcal{M}}$  where  $\mathcal{M}$  is a set containing all the active components.



Now, the importance and meaning of the term  $\lambda E_{ecss}$  or  $\lambda E_{hvb}$  needs to be understood at first in order to define the revised cost structure. The  $u_{hvb}$  is the battery input power and  $\lambda$  represents the EF between chemical energy of fuel and electrical energy. Therefore,  $\lambda \times u_{hvb}$  represents the equivalent fuel cost of using electrical energy per unit time-step. The value of  $\lambda$  varies over time based on the way the HVB charge was replenished in past and will deplete in future. Hence, it is evident that the term  $\lambda \times u_{hvb}$  signifies the equivalent cost of using the electrical power in respect to mechanical power obtained through the fuel consumption of the ICE.

In a multi-mode HEV as presented here, it is imperative to understand that the powertrain can be operated in any of the three modes as described in the chapter 4 at any given time-step, i.e., the vehicle's power demand can be satisfied by either hybrid-electric mode or single-motor EV or even two-motor EV mode. The mode selection is governed by the EMS controller which compares the cost of powertrain operation for three different modes. Hence, there should be a fair comparison among the cost of operation in three different modes in order to select the appropriate mode.  $P_{out}$  refers to power demand at the wheel at any given time-step. The fundamental equation of power balance at the powertrain output is given as follows:

$$P_{out} = P_{ice} + P_{electrical} \quad (5.3.1)$$

, where  $P_{ice}$  and  $P_{electrical}$  are the powers supplied by ICE and electrical machines respectively. Depending on the value of  $P_{ice}$ ,  $P_{electrical}$  can be positive, negative or zero.

If  $P_{electrical}$  is zero, the entire power demand of the vehicle will be sufficed by ICE and since ICE has significantly low efficiency, the loss term for the ICE will be quite

big. In contrast, if  $P_{ice}$  would have been zero, EMGs would have sufficed the entire power demand of the vehicle and since their efficiencies are quite higher than ICE, naturally the loss term would have been significantly small in that case. Therefore, this would lead to an unfair comparison, resulting in activation of either single-motor EV or two-motor EV mode for the entire driving mission. This will obviously violate the charge sustainability criteria of the HVB.

A method to circumvent this biased comparison is presented in [144] where  $\lambda$  is kept as a constant real number. But keeping a constant value of  $\lambda$  does not necessarily ensure the charge sustainability criteria and does not have any physical interpretation except for being a penalty term. One can also think about making the  $\lambda$  adaptable with many available adaptation rules for the EF in the ECMS algorithm. However, this also does not solve issue at hand, rather just give a little robustness to the system to possibly make the system charge sustaining.

A better approach to think about this problem is to understand the underlying meaning of the cost structure and redefine it in a better way. It is very important to understand the underlying meaning of the term  $\lambda$  in order to posit another plausible approach to circumvent the aforementioned problem of unfair cost comparison. Now, it is undeniably true that ultimately all the energy in an HEV must come from the ICE, since there is no other outside source of energy input to the powertrain system. Therefore, whatever the amount of electrical energy is being consumed by the vehicle at a given time-step, ought to be provided by the ICE. Practically, the energy provided by the HVB had been charged by the ICE in the past or will be charged in future. Hence, fundamentally thinking, it can be realized that there is actually ICE loss associated with utilizing the electrical energy from the HVB and that should be

accounted in the cost structure. Accounting for that loss in the cost structure will eradicate the need of including the term representing equivalent fuel cost ( $\lambda \times u_{hvb}$ ) in the instantaneous cost structure. So, the new cost structure can be defined as follows:

$$C_t = \sum_{m \in \mathcal{M}} L_m + \zeta(L_{hvb} + y_{hvb}) \frac{(1 - \eta_{avg})}{\eta_{avg}} \quad (5.3.2)$$

$$\zeta = \begin{cases} 1 & \text{if } y_{hvb} \geq 0 \\ 0 & \text{if } y_{hvb} < 0 \end{cases}$$

The first and second components of the new instantaneous cost structure in Eq.5.3.2 can be referred as objective loss and subjective loss respectively.  $L_{hvb}$  is the amount of loss that occurs across the HVB while delivering  $y_{hvb}$  amount of energy at a given time-step. Hence, the amount of energy that would have been recharged by the ICE can be obtained by adding  $L_{hvb} + y_{batt}$ .  $\eta_{avg}$  is the average efficiency of the path through which the  $L_{hvb} + y_{hvb}$  amount of energy would have been recharged or will be recharged in future. Therefore, the whole second term in Eq.5.3.2, i.e.,  $(L_{hvb} + y_{hvb}) \frac{(1 - \eta_{avg})}{\eta_{avg}}$ , represents the amount of energy lost in the system while recharging the HVB and making the battery capable of emanating positive energy of  $y_{batt}$  amount.  $\zeta$  is a constant integer which can be zero or one depending upon the sign of  $y_{hvb}$ .  $\zeta$  is equal to zero when HVB power is negative ( $y_{hvb} < 0$ ). It is understandable that there is no need of accounting battery is being recharge and not discharged, therefore, there is no loss associated to it. The amount of loss happening in battery and engine while recharging is being accounted in their respective loss terms.

---

**Algorithm 2:** Loss Minimization Strategy for Multi-mode Electrified Powertrain
 

---

**Input :**  $\tau_{rqst_{drvr}}, SOC_{prev}, \omega_{veh_{prev}}, \omega_{ice_{prev}}, \omega_{mot_{prev}}, \omega_{gen_{prev}}$  &  $\omega_{rqst_{drvr}}$   
**Output :**  $U(t)$ , where  $U = \{mode_{optim.}, \Psi_{mode_{optim.}}\}$ , where  $mode = \{1, 2, 3\}$ ,  
 $\Psi_1 = \{\}, \Psi_2 = \{\tau_{ice_{cmd}}, \omega_{gen_{cmd}}\}^T, \Psi_3 = \{\tau_{mot_{cmd}}\}$ ,  
 where  $1 \equiv EV\#1, 2 \equiv HEV, 3 \equiv EV\#2$   
**Initialization :**  $\tau_{out} = \tau_{rqst_{drvr}},$   
 $\omega_{out} = \omega_{rqst_{drvr}}, SOC = SOC_{prev}; Volt = Volt_{prev}$   
**for**  $mode = i, i \in \{1, 2, 3\}$  **do**  
   sweep through all feasible combinations of  
   control variables in mode#i ( $\Psi_i$ )  
   **for** each feasible combinations of  $\Psi_i$  **do**  
     **compute:**  $\omega_{ice}, \tau_{ice}, \omega_{gen}, \tau_{gen}, \omega_{mot},$   
     and  $\tau_{mot}$  corresponding to  $\Psi_i$   
     (use kinematic relationship for the given mode  
     , Eq.2.3.4 for mode#1 & mode#2 and Eq.2.3.5  
     for mode#3)  
     **compute:**  $L_{ice}, L_{gen}, L_{mot}$   
     **compute:**  $y_{hvb} = \sum_{p \in \{gen, mot\}} (y_p + L_p)$   
     **compute:**  $L_{hvb} = f(y_{hvb})$   
     **compute:**  $L_{equiv} = \zeta(L_{hvb} + y_{hvb}) \frac{(1-\eta_{avg})}{\eta_{avg}}$   
     **compute:**  $L_{total} = \sum_m (L_m) + L_{equiv} + L_{auxi}$   
    $\Psi_{i, optim.} = \arg \min_{\Psi_i} \|(L_{total})_i\|$   
 $mode_{optim.} = \min(L_{mode\#1}, L_{mode\#2}, L_{mode\#3})$   
 $U(t) = \{mode_{optim.}, \Psi_{mode_{optim.}}\}$

---

The primary stages of mode and associated control variable selection process are delineated in an algorithmic way in Algo.2. Since the loss minimization is executed numerically, brute force method is employed in HEV and dual-motor EV modes to search the control variable(s) corresponding to minimum loss. The brute force method numerically evaluates the system loss for every combinations of control variable(s) and finds out the control combination corresponding to minimum system loss by brute force. However, a fine discretization of the control variable(s), which automatically

increases the computational burden, is preferable for achieving exact global optimal control accurately. At the present online implementation, a coarser discretization, i.e.,  $\tau_{mot_n}$ ,  $n \in \{1, \dots, 15\}$  and  $\tau_{ice_p}$ ,  $p \in \{1, \dots, 15\}$  with  $15 \times 15 = 225$  evaluations of system loss, has been used in the hybrid-electric mode to keep the computational burden within reasonable limit. Similarly, coarser discretization of  $\tau_{mot_n}$ ,  $n \in \{1, \dots, 15\}$  is used in two-motor EV mode. Naturally, a fine discretization with  $30 \times 30$  grid would lead to more accurate identification of the optimal control.

## 5.4 Simulation and Results

In this section, simulation results obtained from the LMS supervisory control strategy will be presented. As mentioned earlier, the LMS is derived from similar concept used by ECMS which is a well recognized and validated online supervisory control strategy for EMS. Therefore, the results obtained with LMS will be compared with results obtained from the ECMS. It is noteworthy to mention that no mode-switching constraints has been applied on the powertrain operating mode behavior while simulating with both LMS and ECMS. This results in frequent and possibly infeasible mode shifts. Optimally managing the frequency of mode shift by including a penalty term in the cost structure has been discussed in the chapter 4.

Furthermore, a constant optimal value of EF has been chosen while simulating the EMS with ECMS. This ensures that the sub-optimal control provided by the ECMS yields performance with close proximity to global optimal results in the range of 1%-5% [17]. The only unknown parameter in the Eq.5.3.2 is the  $\eta_{avg}$ . Estimating its value is of critical importance since LMS is highly sensitive to the value of  $\eta_{avg}$  and it governs whether LMS can satisfy the charge sustainability criteria at the end of a drive

cycle or not. According to its definition,  $\eta_{avg}$  can be clearly implied as the maximum efficiency with which the battery can be recharged when all the power generated by the ICE will be converted directly into electrical power during hybrid-electric mode. Hence,  $\eta_{avg}$  can be written as follows:

$$f(\eta_{eng}, \eta_{gen}, \eta_{mot}, \eta_{batt}) < \eta_{avg} \leq \eta_{eng}^{max} \quad (5.4.1)$$

The lower limit of the  $\eta_{avg}$  is a function of the operating efficiencies of ICE, generator, motor and HVB. It depends upon the path of energy transfer happening from ICE to battery through the power electronics device/s. Similar to the EF in ECMS, it is also observed in LMS that a optimal constant value of the  $\eta_{avg}$  yields the best result and ensures the charge sustainability criteria of the HVB at the end of a drive cycle. Therefore, the only logical and fair way of comparison is possible when both of the control strategies are compared on the basis of their best achieved results. Hence, an optimal constant value of the  $\eta_{avg}$  is found by trial-and-error method for a given drive cycle such that it satisfies the charge sustaining conditions. Three standard drive cycles, i.e., urban dynamometer driving schedule (UDDS), worldwide harmonized light vehicles test cycle (WLTC), and highway fuel economy test cycle (HWFET) are selected in this article for the performance-based comparative study conducted between ECMS and LMS.

Before going into the comparative study between LMS and ECMS, a couple of crucial results are shown in Fig.5.2 and Fig.5.3. It is noteworthy to reiterate that the transmission plant is an inertia-based model which exhibits dynamics better than a steady-state model. Fig.5.2 and Fig.5.3 show the difference between SOC trajectories and fuel consumption profiles obtained from a steady-state model and inertia-based

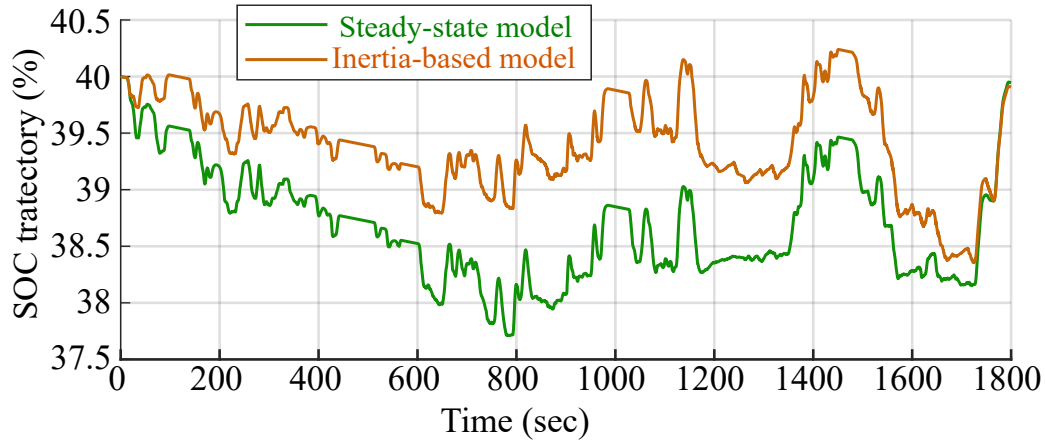


Figure 5.2: Comparison between battery SOC trajectories when inertia-based plant model is used instead of steady-state plant

dynamic model. The comparison is performed keeping the EMS controller same for both of the models.

Fig.5.5(a), Fig.5.9(a), and Fig.5.13(a) present the simulated vehicle speed for UDDS, WLTC, and HWFET respectively. Fig.5.5(b) depicts the SOC traces obtained through ECMS and LMS-based EMS strategies when the vehicle is simulated for UDDS. From Fig.5.5(b), it can be observed that both ECMS and LMS can achieve charge sustenance (final SOC is within 1% deviation of the initial SOC value), given that the optimal value of EF and  $\eta_{avg}$  for ECMS and LMS, respectively are pre-calculated with advance knowledge of the drive cycle. The optimal constant value of  $\eta_{avg}$  and EF found for UDDS are 0.295 (29.5% efficiency) and 2.74 respectively. Fig.5.6, Fig.5.7, and Fig.5.8 depict time-traces of powertrain mode profiles, ICE torques, and ICE speeds, respectively for the two controllers.

In order to have a fair comparison of fuel consumption value between ECMS and LMS, the final SOC value, achieved with both the strategies, should be as close as possible. Fig.5.5(a) depicts the conformity of this requirement and hence, the fuel

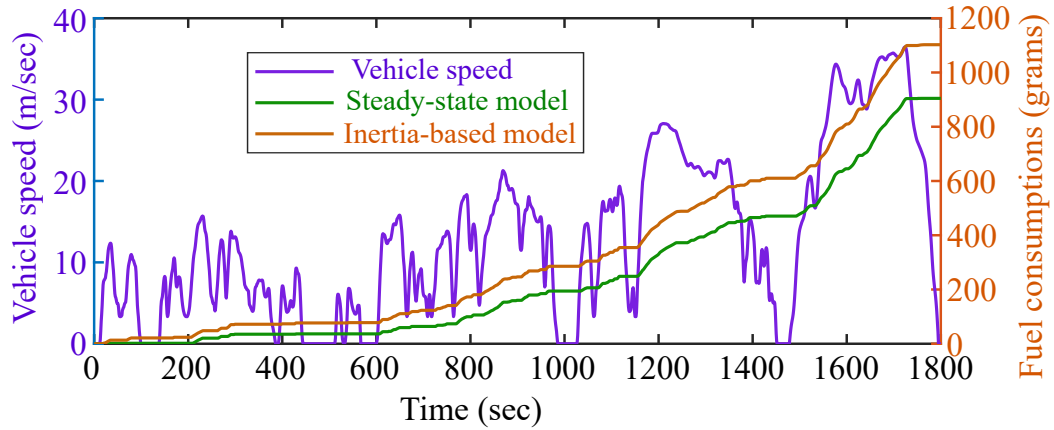
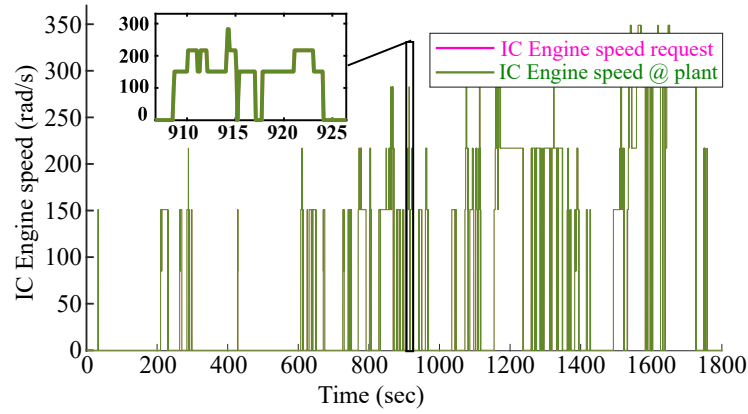


Figure 5.3: Increase of cumulative fuel consumption when an inertia-based plant model is used instead of a steady-state plant for a same drive cycle

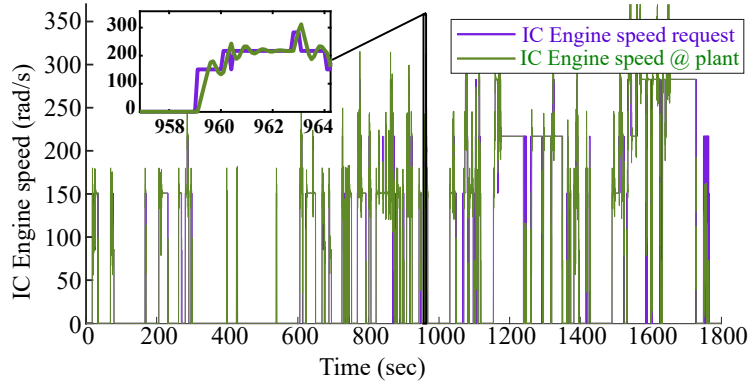
consumption values can be compared. As mentioned earlier, the powertrain suffers from high-frequency and infeasible mode-shift events under both the control strategies due to not imposing any feasibility constraints on the cost-structure for both the strategies. Fig.5.6 corroborates the occurrence of such high-frequency mode-shift events. In order to eradicate the drivability concerns, engendered due to such an impaired behavior of the mode profile, the instantaneous cost-structure should be augmented with a mode-shift penalty in terms of a function of previous time-step's mode. Since the cardinal objective of this article is to validate the prowess of LMS with only focus on minimizing energy loss of the powertrain, drivability improvement is out of the scope of this work. But it will be definitely included in future work.

As depicted in Fig.5.7 and Fig.5.8, the time-traces of ICE operating points in ECMS and LMS nearly follow a similar pattern which results in similar trend in SOC time-traces, as shown in fig. 5.5(b). Tab.5.1 summarizes the comparison of both cumulative fuel consumption and energy loss, which are used as the performance metrics, for all four drive cycles. Similar to UDDS, fair amount of resemblance can be noticed





(a) IC Engine speed response from the steady-state plant w.r.t command from EMS controller

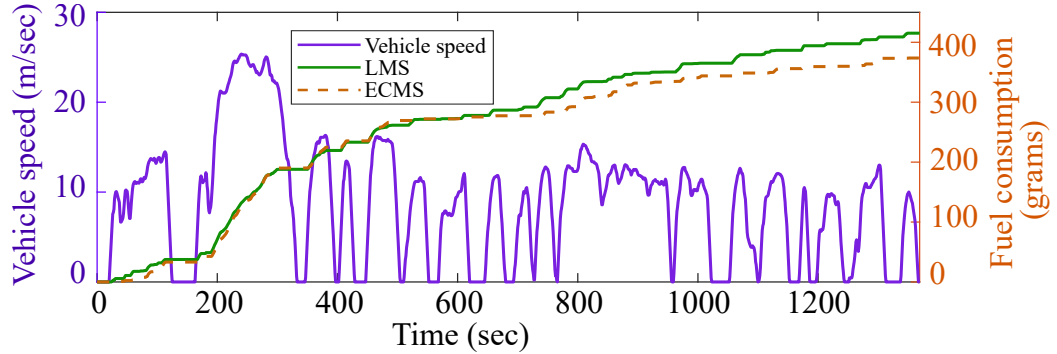


(a) IC Engine speed response from the inertia-based plant w.r.t command from EMS controller

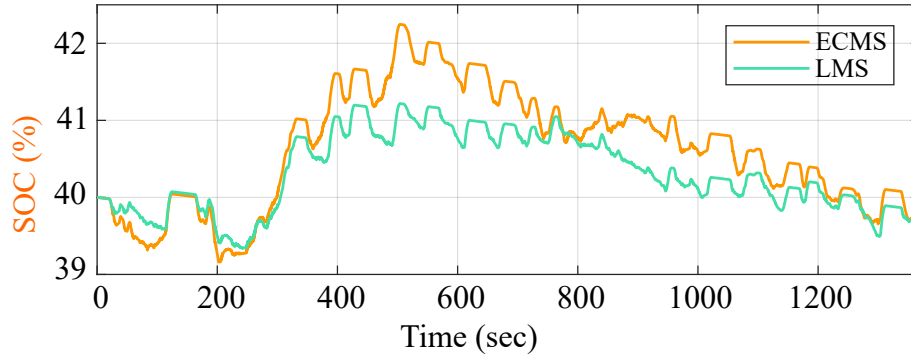
Figure 5.4: Comparison of IC Engine speed response between steady-state and inertia-based plant model

between time-traces of ICE operation with ECMS and LMS for WLTC and HWFET as shown in Fig.5.11, Fig.5.12, Fig.5.15, and Fig.5.16. Effect of such resemblance is clearly visible in the charging and discharging pattern of HVB SOC profiles for both WLTC and HWFET as shown in Fig.5.9(b), and Fig.5.13(b), respectively. The optimal constant value of  $\eta_{avg}$  and EF found for WLTC are 0.3275 (32.75% efficiency) and 2.7195 respectively.

The optimal constant value of  $\eta_{avg}$  and EF found for federal test procedure (FTP)-75



(a) Vehicle's speed &amp; trajectory of fuel consumption



(b) Comparison of battery SOC trajectory for UDDS

Figure 5.5: Vehicle speed, fuel consumption, and SOC trajectories for UDDS

are 0.304 (30.4% efficiency) and 2.7375 respectively. The optimal constant value of  $\eta_{avg}$  and EF found for HWFET are 0.2775 (27.75% efficiency) and 2.7825 respectively. As mentioned earlier, one of the contributions in this article is the inclusion of dynamics of ICE during its cranking and stopping along with the dynamics of other prime-movers attached to PG-set in powertrain modeling. Inclusion of these dynamics results in significant difference in performance of any EMS control strategy compared to the case when powertrain is modeled with steady-state approach and dynamics of both cranking and stopping of ICE are ignored. In steady-state approach, the prime movers, especially ICE operates exactly at the same feasible operating points commanded by the EMS controller. However, this is impossible in reality due to

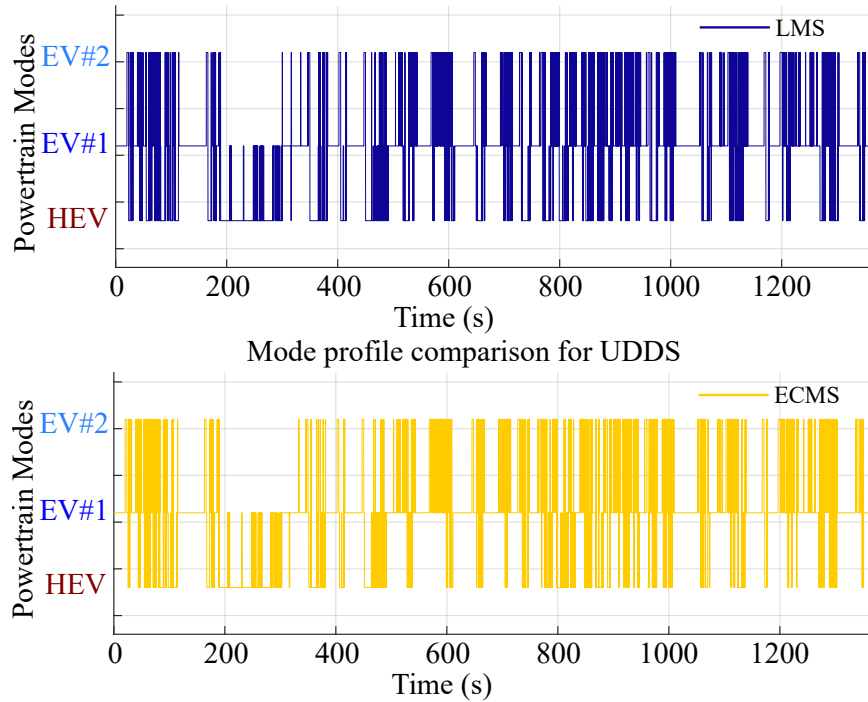


Figure 5.6: Comparison of powertrain operating modes under LMS and ECMS for UDDS

inertia and other physical properties of the components, controller delay, sensor delay etc. Fig.5.20 depicts the speed and torque responses of a low-fidelity ICE model with inertia-based transmission model with respect to controller’s command during hybrid-electric mode. The zoomed views of both speed and torque responses in Fig.5.20(b) and Fig.5.20(c) respectively show that speed of ICE cannot follow the steady speed-request from the controller if ICE receives an oscillating torque-request from the controller. The left-hand-side zoomed window of Fig.5.20(c) shows that torque of ICE model settles down with initial overshoot in response to steady torque requests from the EMS controller. The left-hand-side zoomed window of fig. 5.20(b), which corresponds to left-hand-side zoomed window of 5.20(c), also corroborates the attenuating behavior of the plant ICE-speed in response to a speed and torque requests

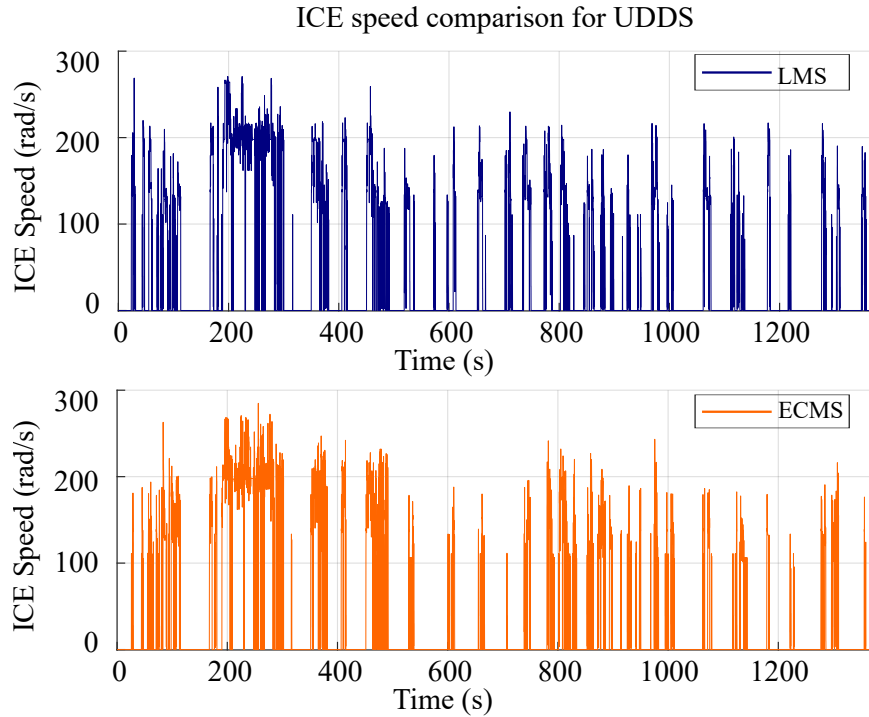


Figure 5.7: Comparison of ICE speed under LMS and ECMS for UDDS

which are steady over a decent amount of time.

Fig. 5.21 depicts the speed dynamics and torque dynamics of ICE during cranking and stopping and also shows the major contribution of generator (EMG#1) during both cranking and stopping. The powertrain always starts with EV#1 mode every time the vehicle launches from zero speed. The generator produces zero torque and rotates in counter-clockwise during EV#1 mode. As soon as the powertrain model receives a command for activation of hybrid-electric mode from the EMS controller, the generator applies positive torque for cranking the ICE, as shown in left-hand-side zoomed window of Fig.5.21(c). As a consequence, the speed of ICE increases in clock-wise and speed of generator rapidly changes from counter-clockwise direction to clock-wise direction due to kinematics of the PG-set, as shown in the left-hand-side

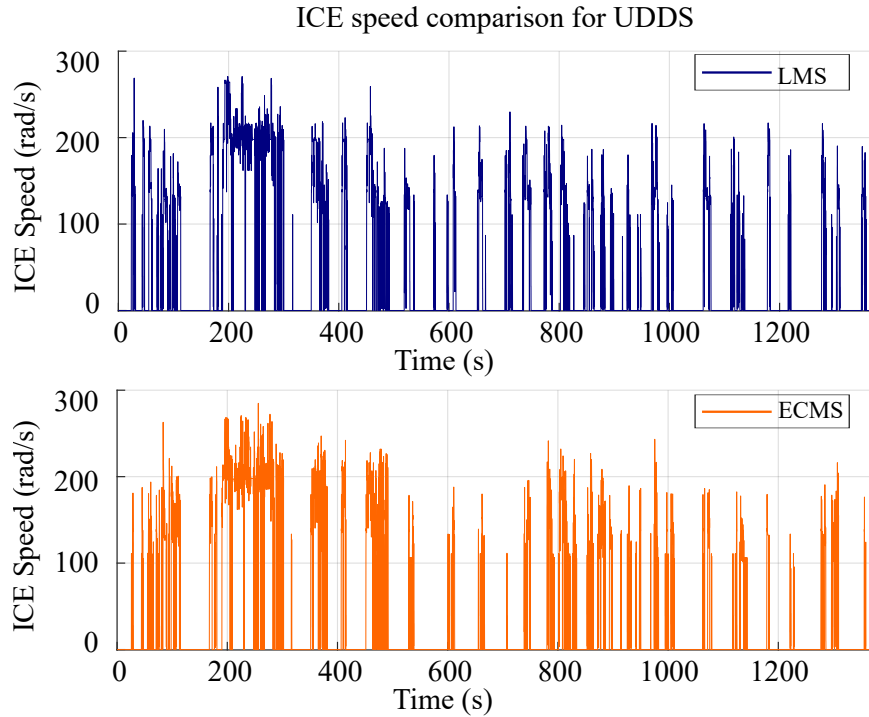
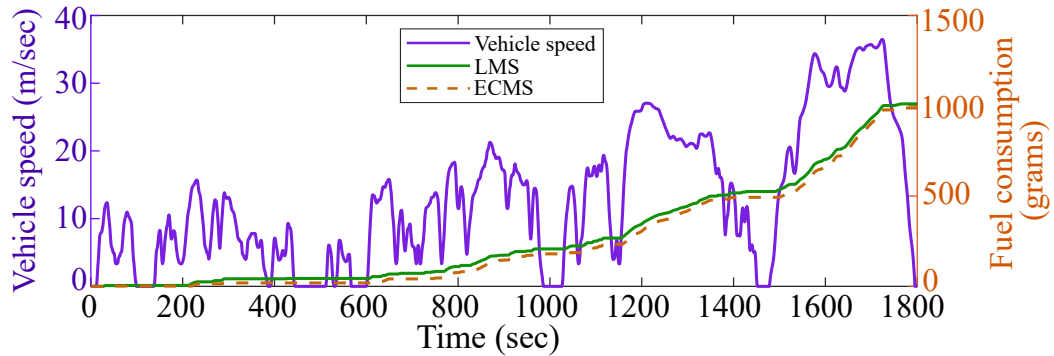


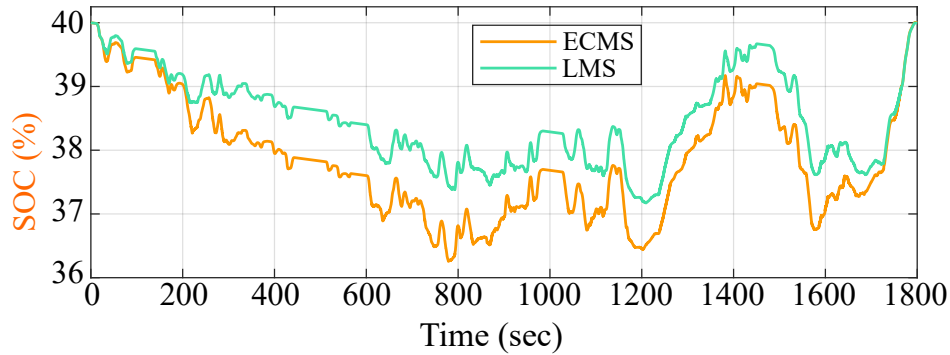
Figure 5.8: Comparison of ICE torque under LMS and ECMS for UDDS

zoomed window of Fig.5.21(b).

Similarly, the generator plays a key role in bringing down the ICE-speed to complete zero when the powertrain receives a mode-shift command from hybrid-electric mode to EV#1 mode. As shown in the right-hand-side zoomed window of Fig.5.21(c), mode-shift from hybrid-electric to EV#1 happens at 265.1 seconds when the ICE torque reduces to zero value due to fuel cut. However, the speed of ICE does not plummet to zero as quickly as the torque does before the mode shifts to EV#1. Hence, rest of the ICE-speed reduction is administered by the generator during EV#1 mode. The right-hand-side zoomed window in Fig.5.21(c) depicts that generator torque gradually increases from a negative value to zero to curtail the ICE speed which is shown in right-hand-side zoomed window of Fig.5.21(b). Proportional controller is employed



(a) Vehicle's speed &amp; trajectory of fuel consumption



(b) Comparison of battery SOC trajectory for WLTC

Figure 5.9: Vehicle speed, fuel consumption, and SOC trajectories for WLTC

to regulate the torque of generator during this curtailment of ICE speed. Finally, the comparison of performance metrics between ECMS and LMS, as furnished in Tab.5.1, corroborates that ECMS performs slightly better than LMS in UDDS, WLTC, and FTP-75, i.e., cumulative energy loss is lesser with ECMS than LMS and consequently cumulative fuel consumption is automatically lesser with ECMS. Both the controllers yield almost same fuel consumption in HWFET.

#### 5.4.1 Implication of Drivability Constraint

Finally, a comparative study is performed between LMS with and without drivability constraint. As discussed in chapter 4, instantaneous optimization strategy assisted

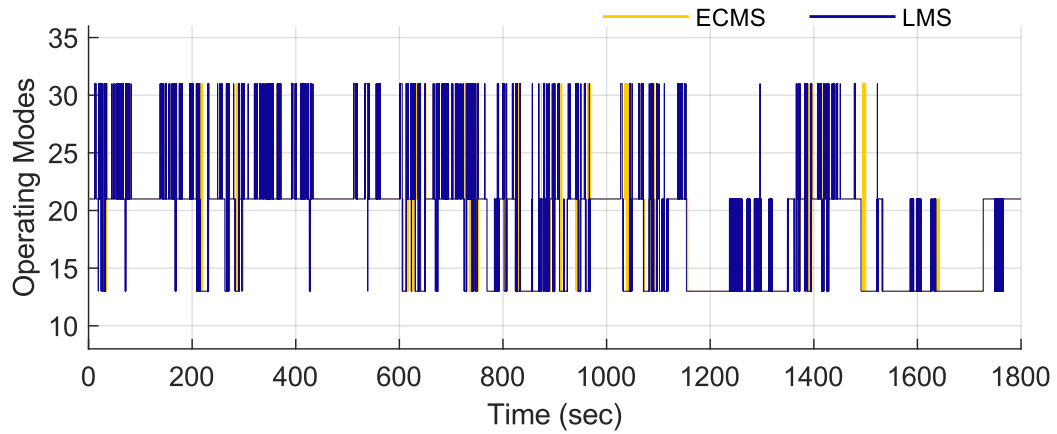


Figure 5.10: Comparison of powertrain modes under LMS and ECMS for WLTC

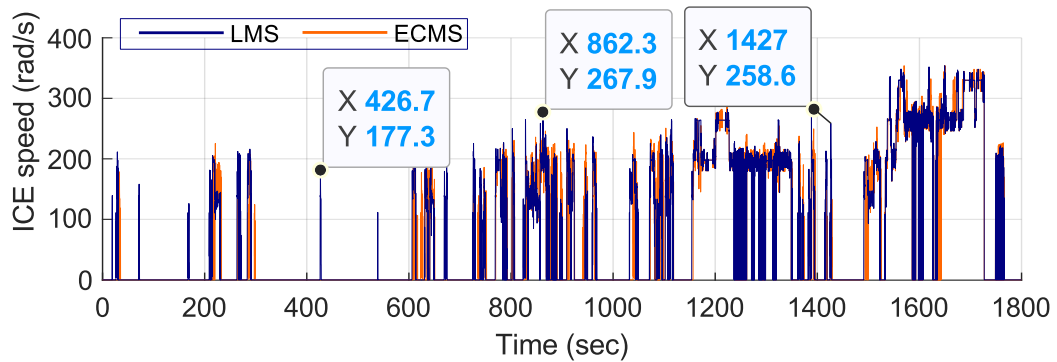


Figure 5.11: Comparison of ICE speed under LMS and ECMS for WLTC

with mode-shift penalty yields feasible mode-shift schedule and gives sufficient time to ICE operation once it is cranked, and prohibits the ICE from cranking immediately after its shutting down. Fig.5.17, Fig.5.18, and Fig.5.19 corroborate the enhanced performance of the LMS when applied on WLTC with drivability constraints. Not only the frequency of mode-shift events has reduced, but fuel consumption has also reduced due to drivability constraint, as shown in Fig.5.19. Hence, it is recommended to include mode-shift feasibility and drivability constraints in the cost function of any instantaneous optimization.

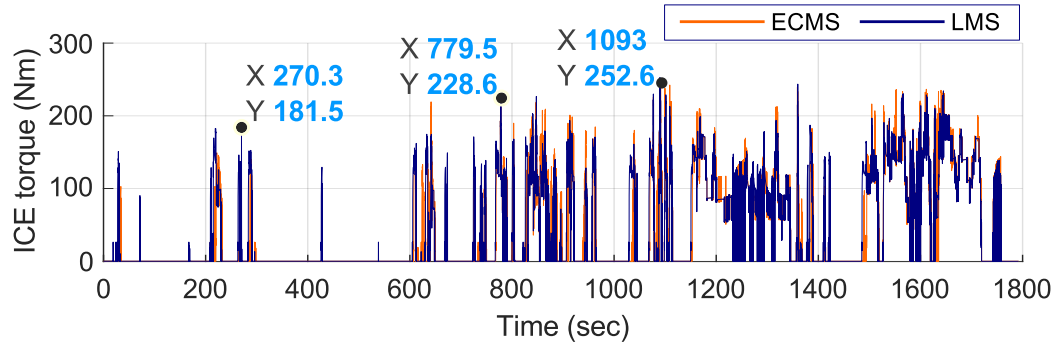


Figure 5.12: Comparison of ICE torque under LMS and ECMS for WLTC

## 5.5 Summary

Based on the minimization of the powertrain’s total system loss, an instantaneous optimal control is proposed for an EMS of a specific electrified powertrain configuration. Theoretical equivalence between ECMS and LMS is elucidated by delineating the underlying concept of the posited LMS built and developed upon a few state-of-the-art online implementations of ECMS. The LMS is applied to a multi-mode EVT having one hybrid-electric mode and two pure-electric modes. The simulation results of LMS are juxtaposed to the optimal ECMS’s results to authenticate the effectiveness of LMS. A few significant findings that have been corroborated in this chapter are enumerated as follows:

- The cumulative system loss over the drive cycle is proportional to the cumulative fuel consumption over the entire drive cycle. Correlation between overall system losses and fuel consumption for four drive cycles firmly corroborate this statement.
- As shown in Tab.5.1, the fuel consumption values obtained through LMS are marginally greater than that of ECMS for all the drive cycles.



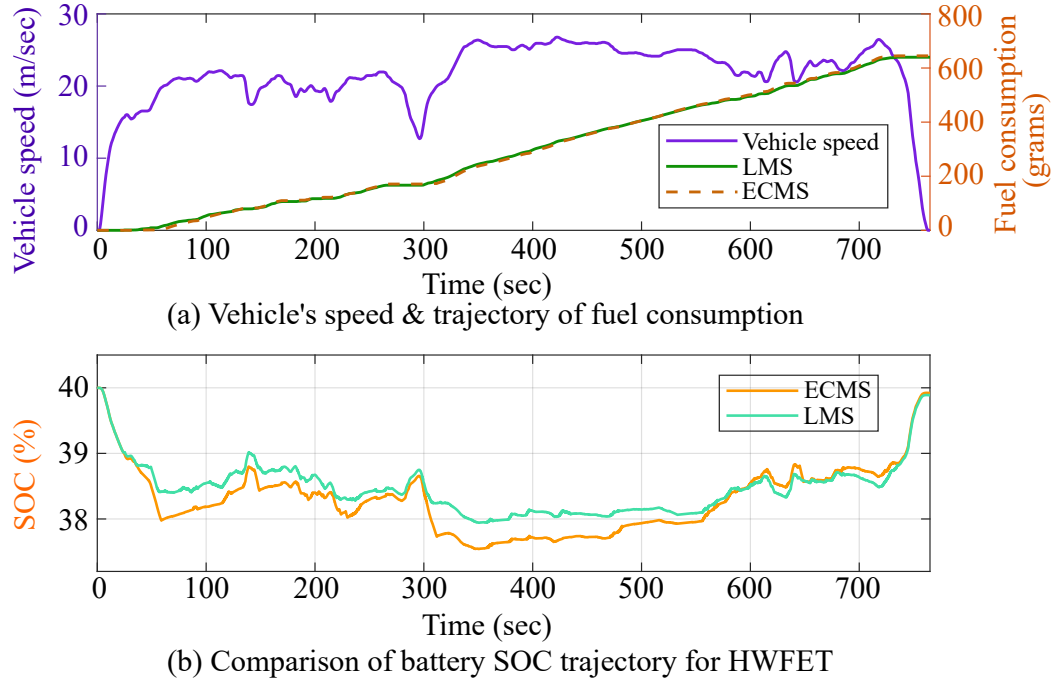


Figure 5.13: Vehicle speed, fuel consumption, and SOC trajectories for HWFET

- The high-fidelity powertrain modeling approach has significant importance in articulation of EMS control strategy. Since dynamics of phenomena like ICE-cranking and ICE-stopping are indirectly related with charging or depleting of the HVB, instantaneous cost-structure should be articulated not only to minimize the system loss and to achieve charge sustenance but also to improve the drivability performance. It should be reiterated that more cranking and stopping phenomena raise the drivability concern and concede more fuel consumption.

The structure of LMS has only one tuning parameter, i.e., average ICE efficiency ( $\eta_{avg}$ ). The simulation results show that the LMS can yield charge sustaining results if the value of  $\eta_{avg}$  is appropriately chosen for a specific drive cycle. It is imperative to understand that the effectiveness of the LMS lies in the meaning of the  $\eta_{avg}$ . The  $\eta_{avg}$ 's

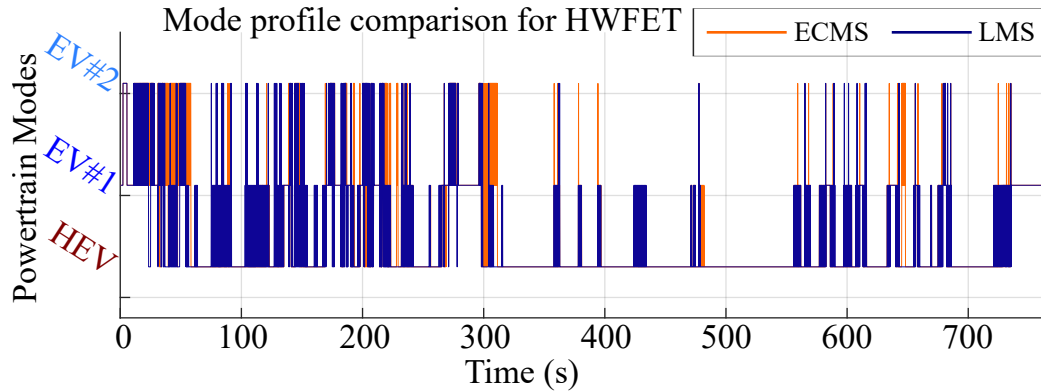


Figure 5.14: Comparison of powertrain mode under LMS and ECMS for HWFET

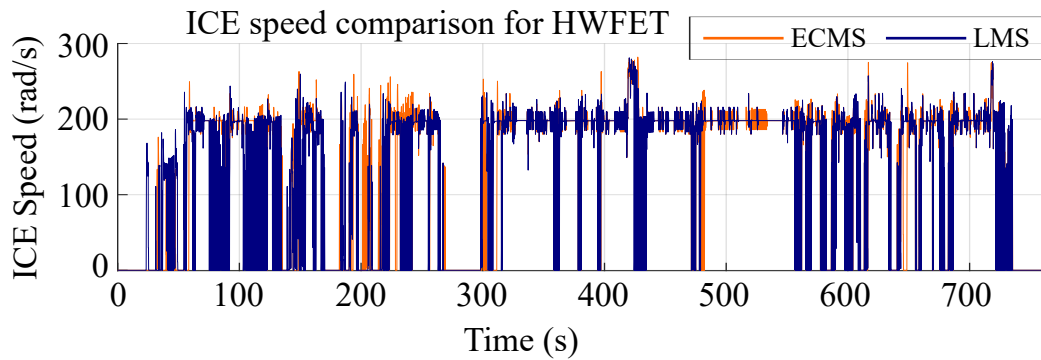


Figure 5.15: Comparison of ICE speed under LMS and ECMS for HWFET

upper limit is fixed depending upon the engine and is neither a function of the drive cycle nor topology or external environment. Furthermore, the  $\eta_{avg}$  has a significant physical interpretation that helps in forming a governing equation for developing an adaptive rule to evaluate the value of  $\eta_{avg}$ . Such adaptive rule can eradicate even the smallest requirement of manual tuning of  $\eta_{avg}$ . Inclusion of adaptation rule for  $\eta_{avg}$  in the instantaneous cost-structure is allotted for the future work. However, the proposed LMS could not outperform the optimal ECMS given that the drive cycle's information is available to both the energy management strategies. This inspires the author of this report continuing the pursuit for an energy management strategy, which can outperform optimal ECMS in real-time simulation.

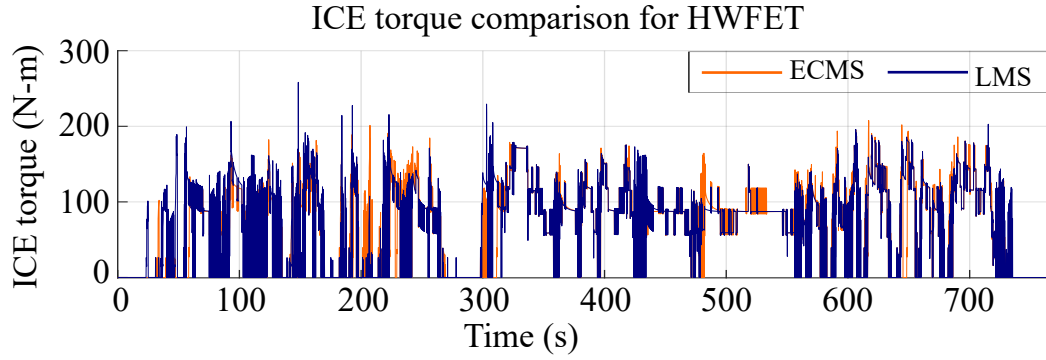


Figure 5.16: Comparison of ICE torque under LMS and ECMS for HWFET

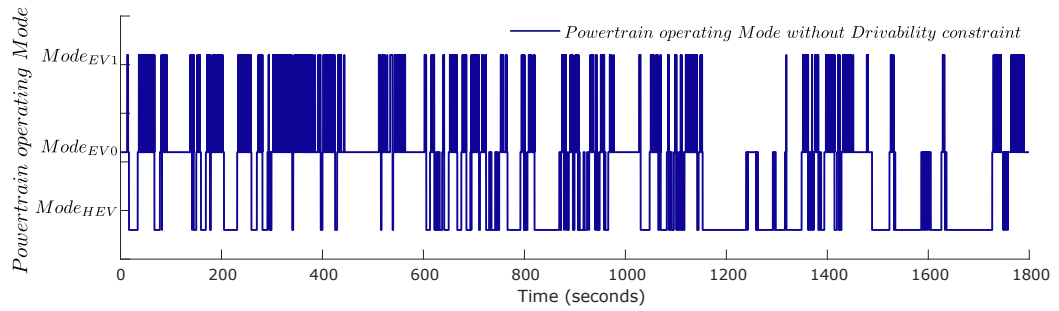


Figure 5.17: Time-series of mode-shift events throughout WLTC under loss minimization strategy without drivability constraint

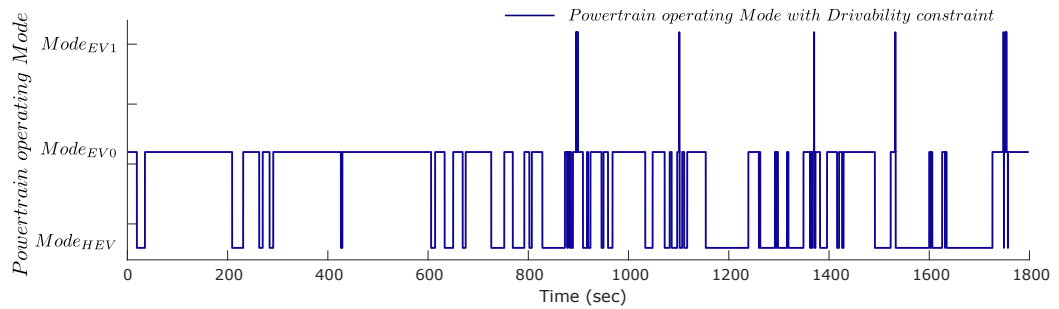


Figure 5.18: Time-series of mode-shift events throughout WLTC under loss minimization strategy with drivability constraint

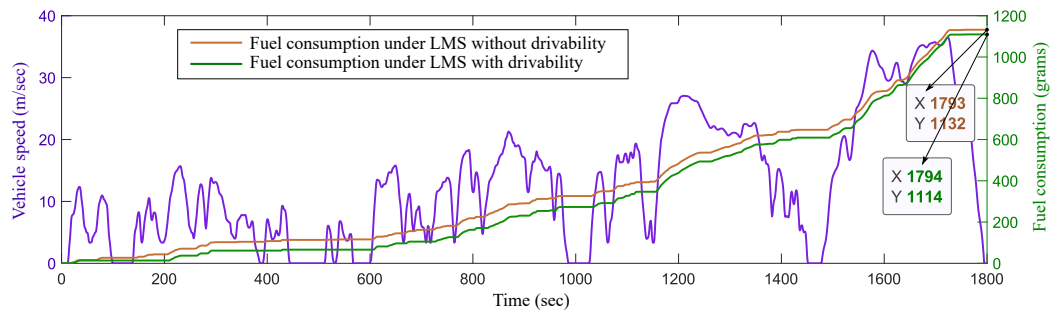


Figure 5.19: Comparison of fuel consumption throughout WLTC under loss minimization strategy with and without drivability constraint

Table 5.1: Comparison of simulation results

Drive cycles	Avg. eff	EF	Fuel consump. (g)		Total system loss (KJ)	
			LMS	ECMS	LMS	ECMS
UDDS	0.295	2.74	415.15	391.98	122510	111180
WLTC	0.3275	2.7195	1203.6	1200.9	354880	335690
FTP-75	0.304	2.7375	806.77	774.01	239130	224940
HWFET	0.2775	2.7825	733.96	716.27	215010	207110

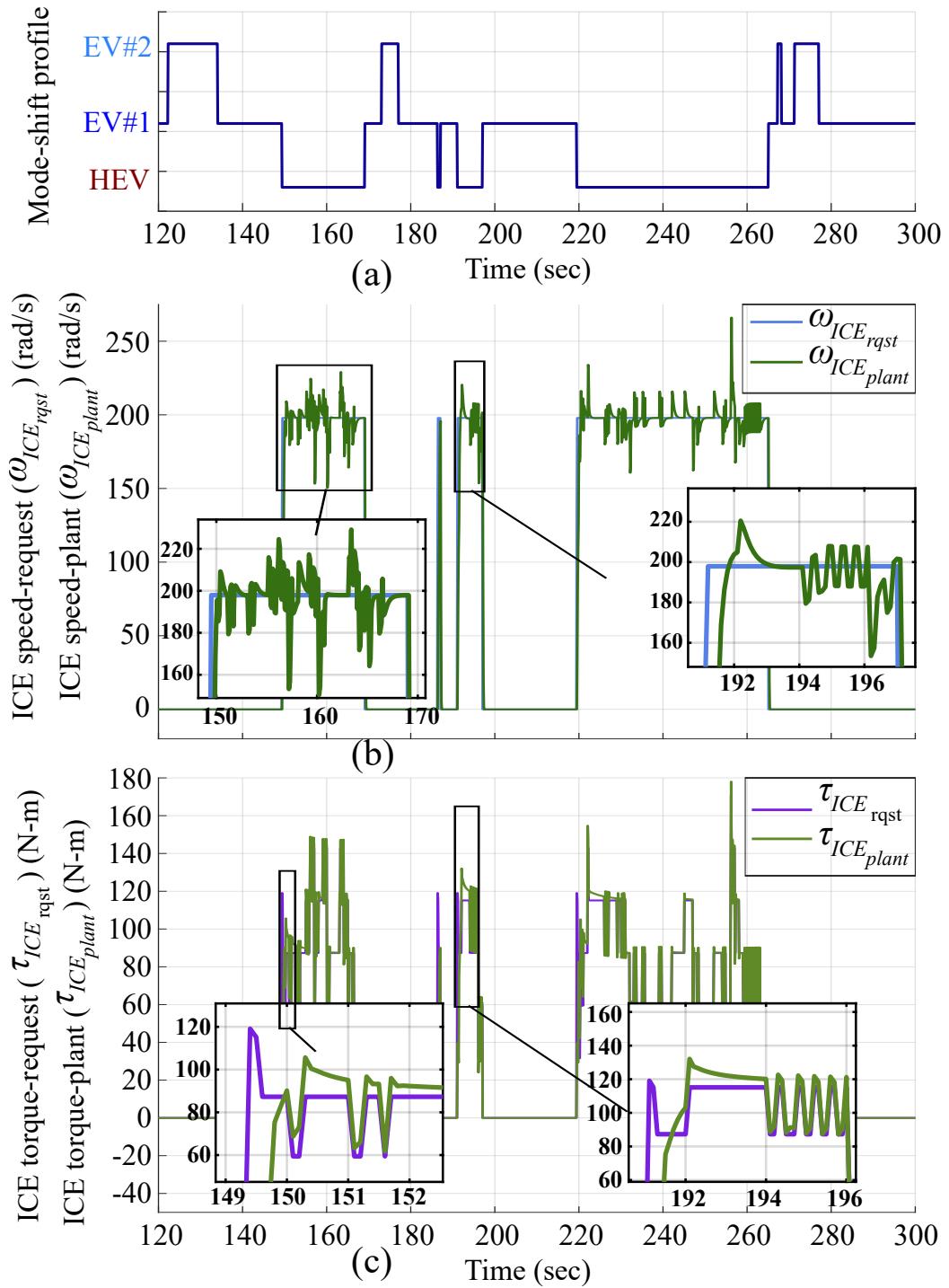


Figure 5.20: Torque and speed responses of high-fidelity ICE w.r.t controller's command

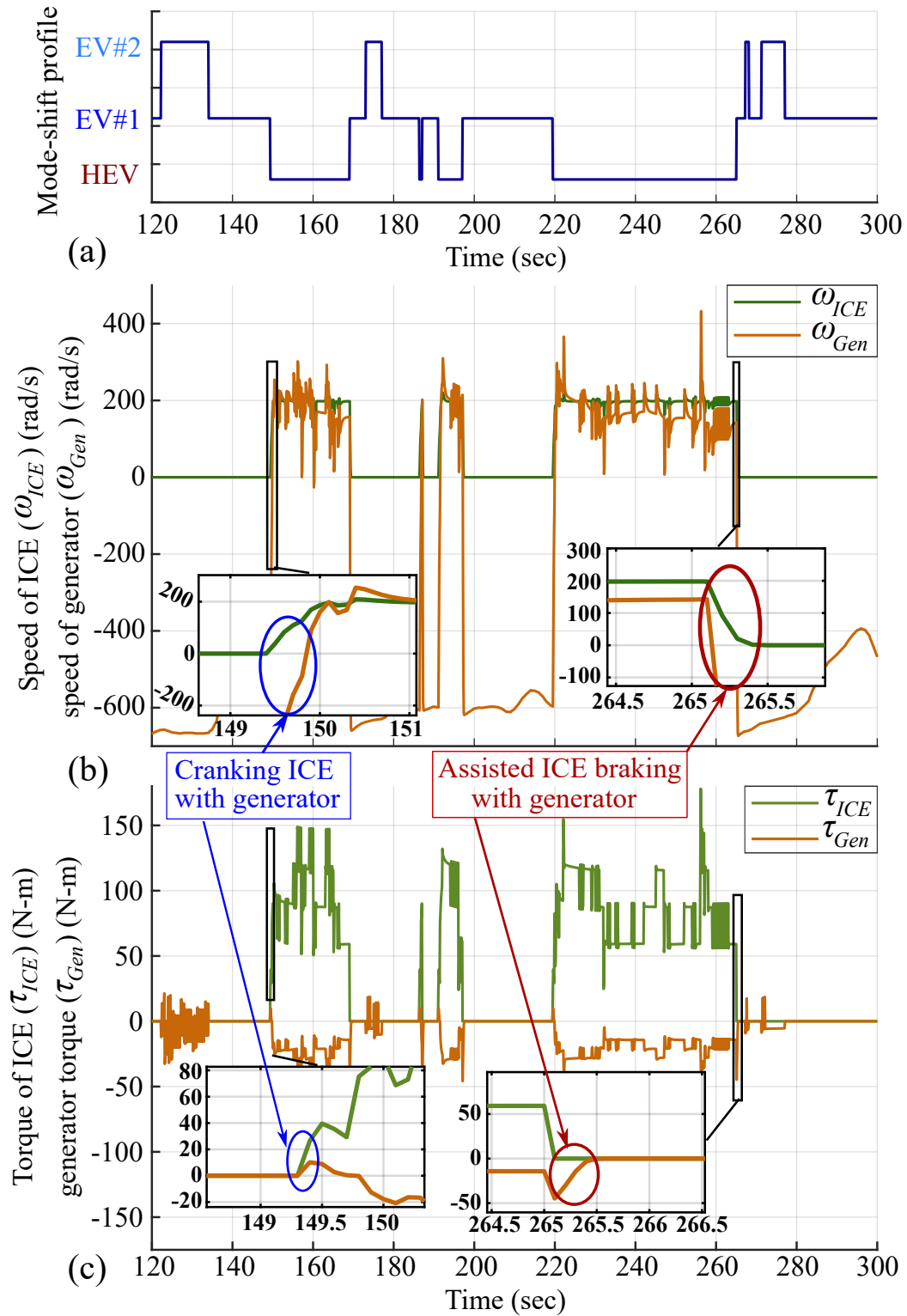


Figure 5.21: Torque and speed dynamics of ICE during cranking and HEV deactivation

## Chapter 6

# Introduction of Reinforcement Learning in Solving the Energy Management Problem of Hybrid Electric Vehicle

## 6.1 Introduction

The previous chapter concluded with the indication of continuing the pursuit for a control strategy yielding near-global optimal performance in real-world driving situations. On the other hand, reinforcement learning (RL) algorithm has been emerging as a lucrative control strategy for optimizing controllable physical systems' performance in real-time since artificial intelligence's intrusion into the sector of optimal control. The RL algorithm mathematically models an intertwined psychological and neurological behavior, exhibited by humans and a few animals, for continuously optimizing decision making capability based on past experiences to substantiate gain in future [25], [26].

RL has declared its strong presence throughout the last decade in the domain of optimal control since researchers started using deep neural network (DNN) for the application of RL. Deep reinforcement learning (DRL) has been successfully applied in several sectors such as robotic control [145, 146], traffic management [147], space exploration rovers [148], and autonomous vehicles [149]. Needless to say that the application of RL in the EMS of the electrified powertrain (ePT) is inevitable, and it is just a matter of time to deploy RL in commercial vehicles.

RL-based agents are especially tailored for solving sequential decision making problems, where long-term return is more prioritized over short-term rewards. This chapter will delineate every rudimentary level of RL algorithm along with one of the well-known RL algorithms, i.e., Q-learning's implementation in solving energy management problem for ePT. This chapter will develop a stable foundation for discussing an advanced and state-of-the-art RL algorithm and its application.



## 6.2 Fundamental of RL based Control Strategy

Before going into the application of RL algorithm in optimizing control policy in real-time, background of learning-based control should be delineated.

### 6.2.1 Fundamentals of Markov Decision Making

We take so many decisions in our daily schedule. Some of them are isolated and have just immediate consequences. But, some of them have both immediate and long-term consequences. Some decision making is sequential where present decision is impacted by the past decisions and the future decisions will be impacted by the present decision as shown Fig.6.1. Markov decision process (MDP) is an effective mathematical tool to model the sequential decision making. The primary events of the Markov decision model are discussed as follows:

Fundamentally, an agent interacts with a target environment ( $\mathcal{E}$ ). At a given time-step in either a finite or an infinite horizon episode, the agent observes only those states of the environment ( $S_t$ ) which are indispensable for deciding actions ( $A_t$ ). The agent will observe two events at the next time-step ( $t + 1$ ) due to the actions  $A_t$  imparted on the environment at the current time-step ( $t$ ).

- The agent will receive a reward  $r_t$  from the environment.
- The agent will see the environment ( $\mathcal{E}$ ) making the transition to a new state  $S_{t+1}$ .

For a finite horizon episode problem, as shown in Fig.6.1, the process mentioned above continues until the agent reaches the terminal state  $S_T$  and receives a cumulative return ( $\mathcal{R}_t$ ) at the end of an episode. The return ( $\mathcal{R}_t = \sum_{k=1}^T \gamma^k r_{t+k}$ ) is denoted as

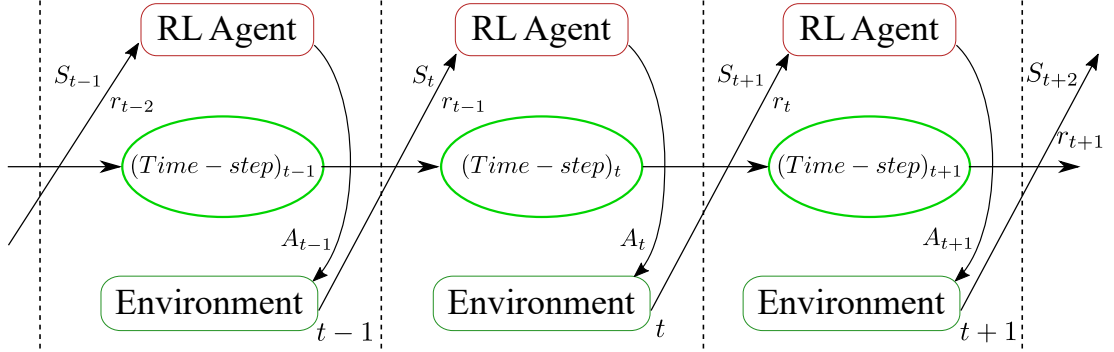


Figure 6.1: Conceptual diagram of sequential decision making, and relationship among state, action, and reward in discrete time-step.

the return caused by the action  $A_t$ . The Return  $\mathcal{R}_t$  can be re-written as following equation:

$$\begin{aligned}
 \mathcal{R}_t &= r_t + \mathcal{R}_{t+1} = r_t + \gamma r_{t+1} + \mathcal{R}_{t+2} \\
 &= r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \mathcal{R}_{t+3} \\
 &= r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^{T-2} r_{T-2} + \mathcal{R}_{T-1}
 \end{aligned} \tag{6.2.1}$$

, where  $\gamma$  is the discount factor ( $\gamma \in (0, 1]$ ) regulating future rewards' contribution to the computation of return from the current time-step.

### States of the Model

A set of environmental information constitute the system state ( $\mathcal{S}$ ) of the decision-making model. The system states provide all the required information to the decision-making agent for choosing a possible action for that state. Markov decision-making model has a n-dimensional state-space  $\mathcal{S} : s^1, s^2, \dots, s^n \mapsto \mathbb{R}^n$ . The environment may possess many other states might not be useful for the decision-making agent.

### Actions of the Model

The set of actions  $\mathcal{A}$  is defined as the finite set  $\mathcal{A} : A_1, A_2, \dots, A_k \mapsto \mathbb{R}^k$  where the size of the action space is  $k$ . Actions are decided by the decision maker to control the system states. The set of actions that can be applied at a state  $s \in \mathcal{S}$  is denoted by  $A(s)$ , where  $A(s) \subseteq \mathcal{A}$ . All the actions might not be applicable to all the states but the set of actions corresponding to each state of the system fall within the set  $\mathcal{A}$ .

### Transition Probability Function (TPF)

Upon application of an action  $A_t \in \mathcal{A}$  in a state  $S_t \in \mathcal{S}$ , the systems makes a *transition* from  $S_t$  to a new state  $S_{t+1} \in \mathcal{S}$ , based on a probability distribution over the set of possible transitions. The transition probability function is defined as  $TPF : S \times A \times S \rightarrow [0, 1]$ , i.e. the probability of ending up in state  $S_{t+1}$  after making control action  $A_t$  in state  $S_t$  is denoted by  $TPF(S_t, A_t, S_{t+1})$ . For all actions  $a$ , and all states  $s$  and  $s'$ ,  $TPF$  should satisfy this relation  $0 \leq TPF(s, a, s') \leq 1$ . Furthermore, for all states  $s$  and all actions  $A(s)$  at state  $s$ ,  $\sum_{s' \in \mathcal{S}} TPF(s, a, s') = 1$ . But for all those actions which are not applicable at state  $s$ ,  $TPF(s, a, s') = 0$ , if system is transitioning from  $s$  to  $s'$ . TPF can be expressed as a conditional probability of landing in state  $S_{t+1}$  at time  $t + 1$  if action  $A_t$  was selected at state  $S_t$  at time  $t$  as  $TPF(S_t, A_t, S_{t+1}) = Prob(S_{t+1}|S_t, A_t)$ . A transition probability matrix (TPM) stores all TPFs for a specific Markov decision making problem.

### Reward Function or Cost Function

As an immediate consequence of the control action made by the decision-making agent, the agent receives an immediate consequence from the environment. Immediate

consequence can be positive in the case of reward and can be negative in the case of cost. If it is a reward, the primary objective of the Markov decision problem would be to maximize the return  $\mathcal{R}$ . On a contrary, if it is a cost, the primary objective would be the minimization of the return  $\mathcal{R}$ . Cost or Reward ( $r$ ) can be mathematically defined either as the immediate consequence of making an action  $A_t$  in the state  $S_t$  and landing in a state  $S_{t+1}$ , i.e.,  $r : S \times A \times S$  or as immediate consequence of making an action  $A_t$  in state  $S_t$ ,  $r : S \times A$ . In this research the second definition will be used as the immediate consequence function. The TPM and the reward function characterize the model of the Markov decision making problem, and their availability will determine what kind of agent or algorithm will be required to solve the Markov decision making problem. An MDP is mathematically represented as a tuple  $\langle \mathcal{S}, \mathcal{A}, TPM, \mathcal{R} \rangle$ .

## Policies

Given an MDP  $\langle \langle \mathcal{S}, \mathcal{A}, TPM, \mathcal{R} \rangle \rangle$ , the rule which governs the agent's choice of action ( $A$ ) for the observed state ( $S$ ) is known as the policy ( $\pi$ ). Policy is a mapping from n-dimensional state space  $\mathcal{S} : S^1, S^2, \dots, S^n \mapsto \mathbb{R}^n$  to k-dimensional action space  $\mathcal{A} : A^1, A^2, \dots, A^k \mapsto \mathbb{R}^k$ . A policy can be deterministic defined as  $\pi : S \rightarrow A$  or it can be stochastic defined as  $\pi : S \times A \rightarrow [0, 1]$ . Policy is the part of the algorithm or agent used for decision making, and its goal is to maximize or minimize the return of the Markov decision problem.

### 6.2.2 Criteria for Optimal Policy in Markov Decision Process

The primary objective of the agent is to find an optimal policy for the MDP through some learning algorithm. If the policy maker is concerned with immediate reward or cost, a simple optimality criterion  $\mathbb{E}[r_t]$  needs to be optimized. If the policy is concerned with immediate as well as long-term consequence, optimality criterion needs to include future consequence with either of these two models,  $\mathbb{E}\left[\sum_{t=0}^h r_t\right]$  and  $\mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t r_t\right]$ .

The first model simply takes a finite horizon of length  $h$  and the agent should optimize the expected rewards or costs over this horizon. The second one is the infinite-horizon model takes all future rewards or costs into account, but the rewards or costs that are received in future are discounted according to how far away in time they will be received. The learning algorithm depends on what optimality criterion is selected.

### 6.2.3 Value Functions

Value functions are essential aspects of MDP because the agent's algorithm computes the optimal policy for the MDP by learning value functions. A state value function helps the algorithm estimate how good it is for the agent to be in a specific state, and an action-value function tells how good to perform a specific action at a particular state. A state  $s_t$ 's value function under policy  $\pi$ , denoted as  $V^\pi(s)$ , is the expected return at the end of the task when starting from state  $s$  and following the policy  $\pi$  after that.

$$V^\pi(s) = \mathbb{E}_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t = s \right\} \quad (6.2.2)$$

Similarly, action-value function, denoted as  $Q^\pi(s, a) : S \times A \rightarrow Q$ , is the expected

return at the end of the task when starting in state  $s_t$ , taking action  $a$  and thereafter following policy  $\pi$ .

$$Q^\pi(s, a) = \mathbb{E}_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t = s, a_t = a \right\} \quad (6.2.3)$$

The value functions satisfy certain recursive property:

$$\begin{aligned} V^\pi(s) &= \mathbb{E}_\pi \left\{ r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots \mid s_t = s \right\} \\ &= \mathbb{E}_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t = s \right\} \\ &= \sum_{s'} TPF(s, \pi(s), s') \left\{ R(s, a, s') + \gamma V^\pi(s') \right\} \end{aligned} \quad (6.2.4)$$

It implies that a state value function is defined in terms of immediate reward and value function of the possible next states weighted by their transition probabilities and discounted by a factor. The value functions play a key role in finding optimal policy because the policy that maximizes value functions for all states (maximize for reward and minimize for cost) is the optimal policy. An optimal policy is denoted by  $\pi^*$ , such that  $V^{\pi^*}(s) \geq V^\pi(s)$  for all  $s \in \mathcal{S}$ . Now the question is how to find the optimal policy? There are two major classes of algorithms to search for the optimal policy of an MDP.

#### 6.2.4 Model-based Algorithm and Model-free Algorithm

Based on MDP's characteristic model's availability, i.e., TPM and immediate consequence function, two types of algorithms can be employed to search for the optimal policy. Model-based algorithms are employed when the agent is familiar with both

TPM and immediate consequence function beforehand. Model-based algorithms exist under the general name of dynamic programming (DP). On the other hand, model-free algorithms, under RL's name, do not rely on the availability of TPM of the MDP. Both model-based and model-free algorithms necessitate immediate consequence function. Model-free algorithms rely on the interactions, i.e., real-time interactions or simulated, of the agent with the environment. The agent must explore the MDP since it does not have the exact model of the MDP.

### 6.2.5 Generalized Policy Iteration

For both model-based and model-free algorithms, an iterative process must be employed to evaluate and improve the current policy. Starting with a naive policy, either of DP or RL algorithm will evaluate and improve the policy periodically to reach the optimal policy, as shown in Fig. 6.2.

#### Policy Evaluation

The policy evaluation step estimates the present policy's usefulness by calculating state-value functions for all states  $s \in \mathcal{S}$ . In model-based algorithms, TPM and immediate consequence function are used to compute the value functions. In model-free algorithms, actions can be taken according to the environment's policy, and response is recorded in every sampling time. Value functions are estimated from these recorded responses from the environment.

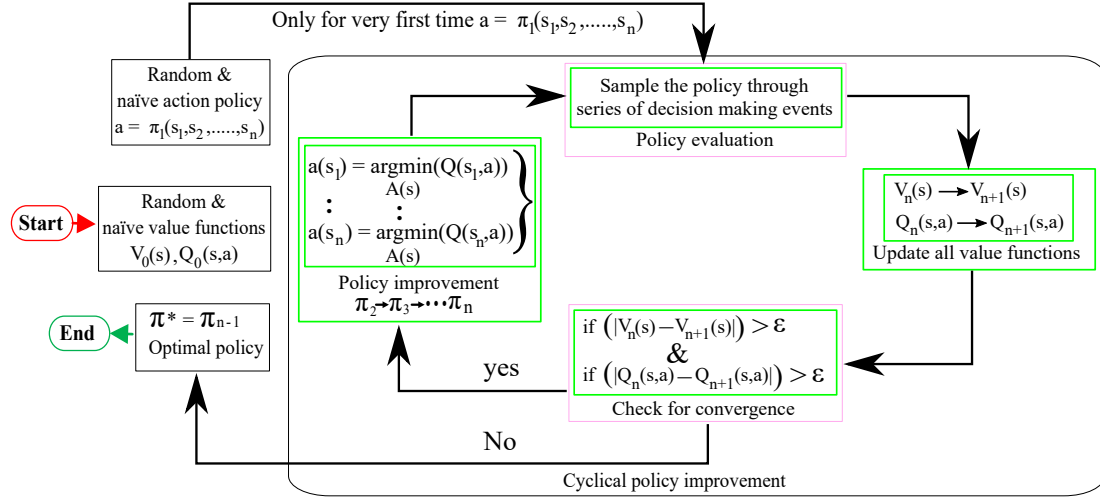


Figure 6.2: Schematic representation of sequential decision problem.

### Policy Improvement

The action-value functions are evaluated for every state to find any better possible actions that can better action-value function than the current action-value function. This step computes an improved policy  $\pi'$  from the current policy  $\pi$  using the information in  $V^\pi$ . Since the model-free algorithm is best suited for most real-life decision-making processes, the next focus will be on the RL algorithm.

### 6.2.6 Reinforcement Learning: Model-Free Algorithm to Obtain Optimal Policy

As stated earlier, the RL algorithm is used to find the optimal policy when the TPM of MDP is unavailable to the agent. The unavailability of the exact MDP model compels the RL to sample MDP to accumulate estimated stochastic information of the MDP. After gathering enough data by sampling, the value functions are calculated, and then the optimal policy is found as it is found in model-based techniques. This



approach takes longer because it first learns the transition probabilities and immediate consequence function from interaction with the environment, then estimates the value functions, and finally finds the optimal policy by policy iteration. This approach is called indirect RL. Another approach directly starts estimating the value functions, even before estimating the proper model of MDP. This approach is called indirect RL. A combination of these two approaches can accelerate the learning process.

### Temporal Difference Learning

Another thing must be decided: estimate the value function without even estimating the MDP's model. Should we wait till the end of an episode to assess the state value function and an action-value function? It is not fair to assign a value function to a state or action after waiting so long. Hence, the estimated value function is modified based on the immediate reward and the estimated value function of the next state. This mechanism is known as temporal difference (TD). Undeniably, TD learning would be the most appropriate algorithm to update the estimated value functions when TPM of the MDP is known apriori. TD learning, which is also known as model-free learning or prediction-based learning, is the cornerstone for the future of artificial intelligence (AI). In simple words, TD learning is the mechanism of updating an estimate with the help of another estimate. The agent selects an action at the current state and obtains an immediate reward and next state from the environment. Based on the estimated value function of this next state and the immediate reward, estimated value function of the current state  $\tilde{V}$  and state-action  $\tilde{Q}$  under policy  $\pi$  is updated. For  $TD(0)$  learning method, only the next state's value function is used to update current state's value function in the way:  $V_{k+1}(s) =$

$V_k(s) + \alpha \left\{ r + \gamma V_k(s') - V_k(s) \right\}$ , where  $\alpha \in [0, 1]$  is the learning rate deciding the rate of update of value functions. The learning rate should be high at the beginning and decrease as time progresses.

### Exploration-Exploitation Ratio

The strategy of choosing an action by the RL agent needs to be decided. Since the RL agent does not have exact knowledge of  $V(s)$  and  $Q(s, a)$  for each state  $s \in \mathcal{S}$  and action  $a \in \mathcal{A}$ , it should not rely on *greedy* policy when selecting an action in a state. The agent should exploit the available information about value functions as well as explore the entire action space. A separate exploration mechanism ensures that the best action (according to current estimates of action value functions) is not chosen always. At the naive stage, the learner should explore more than exploit. As time proceeds, *exploration-exploitation* ratio can be reduced. A  $\epsilon - greedy$  policy instead of the *greedy* policy ensures desired exploration and exploitation. The learner takes the current best action with a probability of  $(1 - \epsilon)$  and chooses action randomly with a probability of  $\epsilon$ .

### Q-Learning Algorithm

One of the popular algorithm for action-value function learning in TD mechanism is Q-learning. Like  $TD(0)$  method, action-value function  $Q(s, a)$  is updated with the following equation:

$$Q_{k+1}(s_t, a_t) = Q_k(s_t, a_t) + \alpha \left\{ r_t + \gamma \max_a Q_k(s_{t+1}, a) - Q_k(s_t, a_t) \right\} \quad (6.2.5)$$



Figure 6.3: Agent-environment interaction for Q-learning.

The Q-learning works in a simple way. At any state ( $S_t$ ) of an episode, the agent first interacts with the environment by taking a feasible action ( $A_t$ ). As a consequence, the environment gives a reward ( $r_t$ ) and takes the agent to next state ( $S_{t+1}$ ), as shown in Fig. 6.3.

The learner chooses an action from the set of actions in the next state based on whichever action gives maximum estimated action-value function  $\arg \max_{a(t+1)} [Q_k(s_{t+1}, a)]$ . Now, the action value function corresponding to current state and action ( $Q_{k+1}(s_t, a_t)$ ) is updated based on previous value ( $Q_k(s_t, a_t)$ ). The schematic of Q-learning is shown in detail in Fig. 6.4. One advantage of Q-learning is that the policy evaluation and policy iteration take place simultaneously.

### 6.3 Application of RL Agent in Real-Time Control for EMSs

There are copious drive cycles in the real-world driving scenario. Optimizing an energy management controller for a specific drive cycle or a group of drive cycles is possible, but all feasible drive cycles are neither possible nor a smart idea. Instead, incorporating a learning technique in the controller is a smart idea. The controller

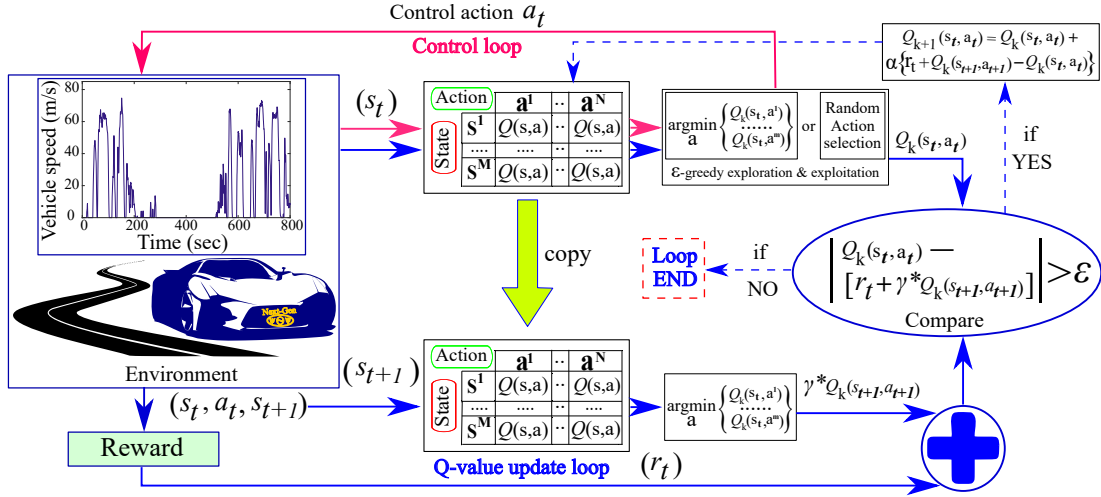


Figure 6.4: Schematic of Q-value function update with continuous policy improvement.

can adapt its control policy according to the most recent driving scenarios with the learning technique's help. RL agents can be deployed as such adaptive controllers that keep updating control policy without knowing future driving situations. EMS's control policy must be articulated so that the policy yields global-optimal control in all real-world driving scenarios. Control of energy management problem for any drive cycle is a perfect example of MDP, where the controller sequentially decides for each sample point of the drive cycle.

### 6.3.1 Justification of Application of RL Agent in Real-Time EMS

In real-world driving situations, where the next driving state is entirely unknown to the controller, the perfect model of MDP is unavailable. If the drive cycle is known as *a priori*, the perfect MDP model (transition probabilities and immediate consequence function) could be articulated, and then the optimal policy for the MDP can be

obtained through DP.

Now the question is, why is a perfect model of MDP unavailable or cannot be articulated for real-world drive cycles with uncertain future driving states? The answer to this question justifies an RL agent’s requirement for energy management in real-time and carrying out this research. MDP is comprised of two main elements, namely, TPM and immediate consequence function. For electrified vehicular applications, instantaneous fuel consumption, instantaneous tailpipe emissions, and instantaneous battery SOC change are a few examples of immediate consequences. The question can be even narrowed down as that why TPM cannot be articulated for real-world driving? As explained earlier, TPM can be expressed as a conditional probability of landing in a state  $s'$  at time  $t + 1$  if action  $a$  was selected at state  $s$  at time  $t$ . For real-world driving, the controller lands on an unfamiliar state  $s'$  when it takes a specific action  $a(s)$  at state  $s$ . Such unfamiliarity happens because the state vector comprises power demand and other powertrain variables like battery SOC, the vehicle’s speed, and acceleration of the vehicle. If the state vector would have only contained powertrain variables, then TPM would have been modeled easily. The vehicle’s power demand for the next time-step is always uncertain, and power demand cannot be eliminated from state variables. Energy management strategies with predictive power demand are found in the literature. However, this research focuses on articulating a control strategy based on uncertain power demand to robust against any unprecedented power demand. Moreover, RL-based control policy’s performance can reach the closest to that of DP-based control among all the existing online control policies. Hence, model-free algorithms or RL-based algorithms are employed to learn and obtain a near-global optimal control policy.

### 6.3.2 Ultimate Objective of Control Policy in terms of Global Optimality and Real-Time Application

Looking at the drawbacks of the existing real-time control strategies, as discussed in the chapter chapter 4 and chapter chapter 5, an ultimate objective can be prescribed, and then it will be discussed that how closer to the ultimate objective can be achieved via RL-based agent.

#### Ultimate Objective

The ultimate objective is to yield global-optimal control for all real-world driving situations. The controller should produce instant global optimal control for every real-world driving situation. The controller should execute optimization within the control time-step like PMP or ECMS control strategy, but the optimization should be global optimization instead of local optimization.

#### Limitations Against Achieving Ultimate Objective

The controller cannot produce global-optimal control for an unfamiliar driving situation at the very first interaction. A specific time-period should be given to the controller to produce global-optimal control for any new driving situation. A few learning algorithms using  $\epsilon - greedy$  policy can take significant time to yield optimal control action for a state that is less encountered by the learner. However, with efficient exploration, the convergence time of the learning algorithm can be reduced. The bottom line is that the controller should take a certain amount of time before making global-optimal action for a new driving situation.

A specific real-world driving situation can be described most simply with a single

state, namely, vehicle speed. The same driving situation can be described by multiple states, namely, speed, acceleration, road grade, battery SOC, tentative distance from destination, wind speed, ambient temperature, and many more. The number of states should be limited to four or five to keep the computational burden within the feasible limit.

### 6.3.3 Approach to Update Control Policy

The control policy should update continuously in real-time to satisfy the objective mentioned above that the policy should yield near-global optimal control for every real-world driving situations. A few plausible mechanisms for updating the policy are as follows:

#### Continuous Updating Policy through Onboard RL

The vehicle can be deployed on the road with an onboard RL agent and an initial policy. The initial policy needs not to be global-optimal or even near-global optimal because the onboard RL algorithm will eventually elevate the policy to near-global optimality as it encounters more drive cycles. The RL agent receives required state signals from sensors and other controllers at each time-step as the vehicle runs on the road. The RL agent learns the action value functions  $Q(s, a)$  continuously in real-time with the Q-learning algorithm and improves the control policy.

As shown in Fig.6.4, the Q-learning only requires the current state information to update the Q-value function of a state-action combination. Whenever the agent interacts with new driving states during this process, it estimates  $Q_{esti}(s_{new}, a)$  for those new states and learns the near-global optimal action-value function  $Q(s_{new}, a)$

for those states by real-world interaction as well as simulated interaction. Q-learning not only converges to near-global optimal value of  $Q(s_{new}, a_t)$ , but it improves the policy at each time-step by choosing an action with estimated optimal Q-value at each time-step. However, it will be better to follow  $\epsilon - greedy$  policy.

**Additional parallel computation is required to expedite the learning process.** If the vehicle sensor's drive-cycle data is received per second, the controller will have much idle time, and the learning process will be slow. The controller also needs many iterations of the current drive cycle to converge to expected Q-values for all state-action combinations. Hence, the controller needs more drive-cycle data both to utilize the idle time and execute many iterations. Nevertheless, how can be drive-cycle data obtained before the real-world drive cycle terminates? **Here comes the Markov chain model (MCM) for continuously generating random drive cycles using current driving data.** An updating MCM can be articulated using the current driving data of the vehicle. The MCM stores the transition probabilities of power demand for the recent past and ongoing driving scenarios. Using the MCM, numerous drive cycles can be generated within one second, and these drive cycles can expedite the agent's learning process. The RL agent with Q-learning and MCM's updating functionalities can be embedded in an electronic control unit (ECU) or an onboard computer to execute the RL agent-based EMS for an electrified powertrain. **This approach of expediting the learning process with an updating MCM will make this research different from existing literature.**



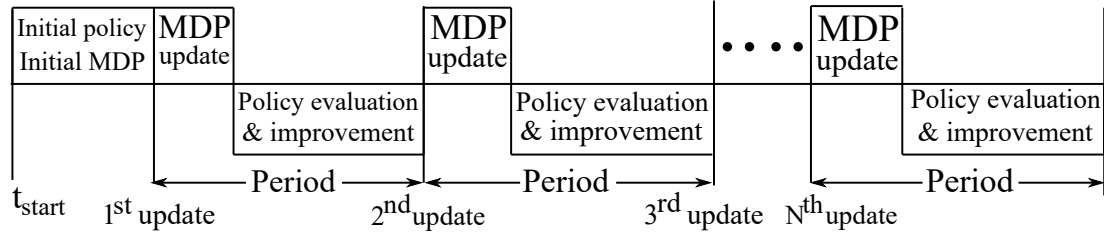


Figure 6.5: Concept of periodical update of control policy with on-board ECU or computer.

### Periodic Updating Policy through On-Board RL

The periodic updating mechanism works like the continuous update, with the only exception is that the update of policy is periodical instead of continuous. The periodic update reduces the intensity of real-time computation to make it feasible for onboard ECU or computer. The RL can learn value functions based on the periodically updated MDP model, and it can be used to update the control policy periodically, as shown in Fig.6.5.

### Periodic Updating Policy through Off-Board RL

This updating mechanism's working principle is the same as the previous one, but there is a difference in the execution. Instead of using an onboard computer, a laboratory-based high-performance computer is used for periodically updating the RL agent's optimal policy. The onboard sensors and ECU can continuously collect required state signals, and an ECU will be responsible for updating the MDP model periodically. The periodically updated model of MDP is then sent to the laboratory via the internet. The updated MDP model is used in the laboratory to update the RL agent's policy, which is sent to the vehicle again through the internet. The whole process is shown in Fig.6.6.

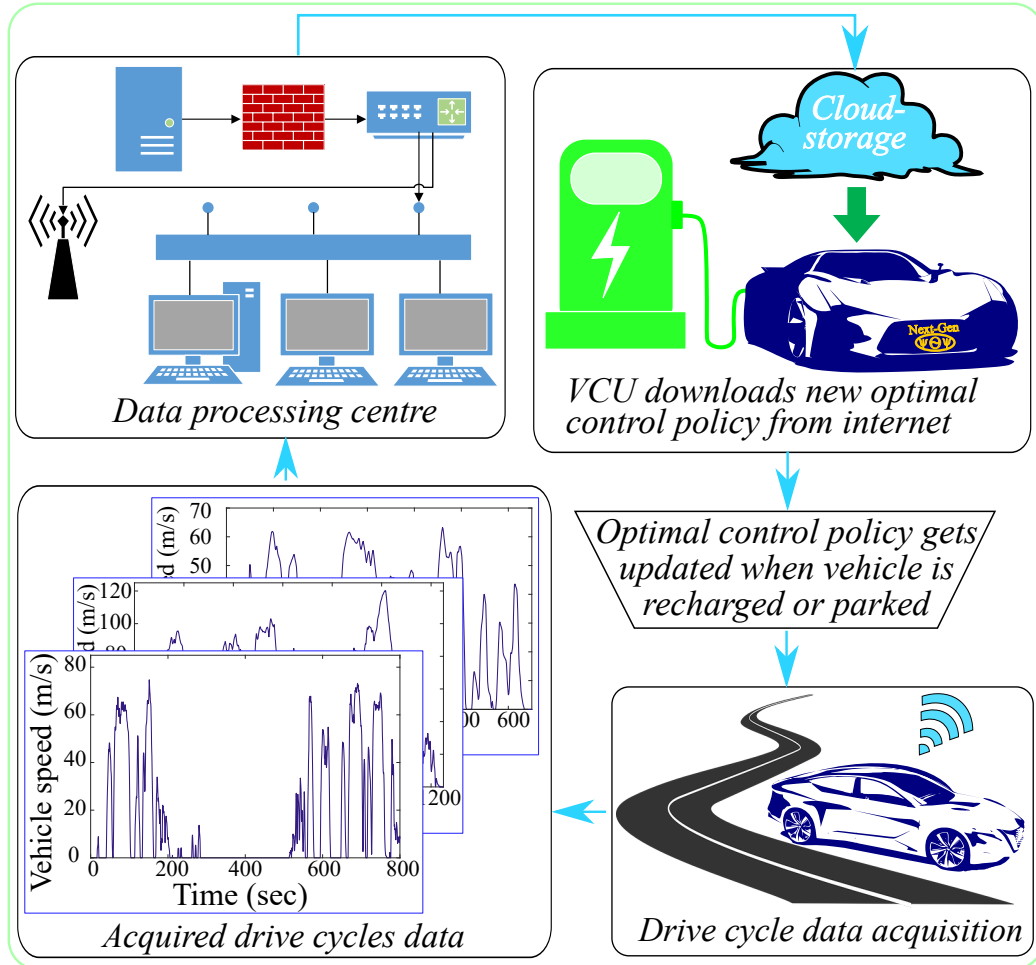


Figure 6.6: Concept of periodic update of control policy with off-board high-performance computer.

### 6.3.4 RL Agent with Q-Learning as Energy management Strategy

The fundamental algorithm of Q-learning with its basic components are delineated in 6.2.6. This subsection will delineate the steps required to implement Q-learning for obtaining an global optimal policy for EMS of a electrified powertrain.



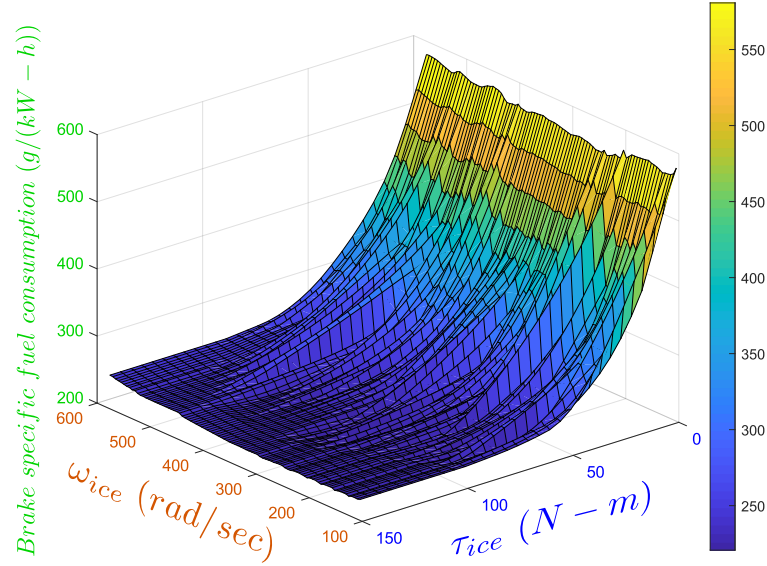


Figure 6.8: Tentative map of brake specific fuel consumption of the Atkinson cycle-based ICE used in Toyota Prius III generation.

Table 6.1: Vehicle specification of Toyota Prius in brief

Name of parameter	Value (unit)	Name of parameter	Value (unit)
Vehicle mass	1392 ( <i>kg</i> )	Vehicle inertia	145.78 ( <i>kg - m<sup>2</sup></i> )
Wheel radius	0.31725 ( <i>m</i> )	Final drive ratio	3.268
RL_Coeff_A	35.53	RL_Coeff_B	0.327
RL_Coeff_C	0.0227	Air density	1.2754 ( <i>kg/m<sup>3</sup></i> )
Height of vehicle	1.75 ( <i>m</i> )	Width of vehicle	1.48 ( <i>m</i> )
Aero drag co-eff	0.325	Front area of vehicle	0.86 × height × width

3rd generation P410 transaxle) has two DOF due to two planetary gear-sets as shown in 6.7, ICE torque and ICE speed are the two variables of the decision vector. The discretized vector of ICE speed and ICE torque is given in 6.2. The specifications of the powertrain’s cardinal components are furnished in Tab.6.1. The brake specific fuel consumption (BSFC) map of an SI-engine with peak torque of 142 N-m and peak power of 73 kWatt is depicted in Fig.

The same discretized vehicle power demand vector, used to generate states for MCM,

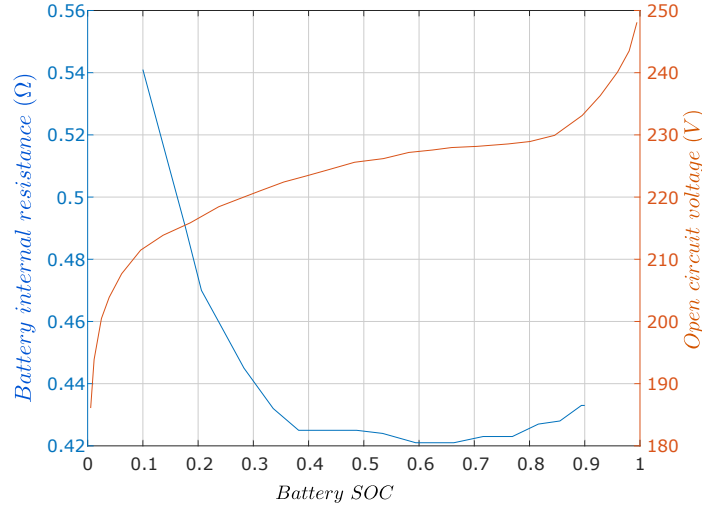


Figure 6.9: Tentative maps of the battery pack’s Open circuit voltage and the battery pack’s internal resistance are expressed as a mathematical function of battery SOC, used in Toyota Prius.

is used here too. Instead of using discretized vector of vehicle’s speed, discretized vector of battery SOC is used here as the second dimension of state of environment.

### Articulation of Immediate Cost table

The next important component of Q-learning implementation is the articulation of immediate consequence function corresponding to different combinations of state and action values. Since energy management is a minimization problem, any consequence will be considered as a **cost**. The immediate cost matrix has a dimension of  $4346 \times 2472$ , producing 2472 actions for each of the 4346 states. There is no fixed rule for articulating an immediate cost matrix, but this articulation plays a vital role in obtaining the global optimal control policy in real-time EMS. A separate MATLAB function ( $Immediate\_Cost(State, Action)$ ) has been created to calculate the immediate cost associated with every feasible combination of state and action. The algorithm used in  $Immediate\_Cost(State, Action)$  is shown in Fig.6.12. If all the column values

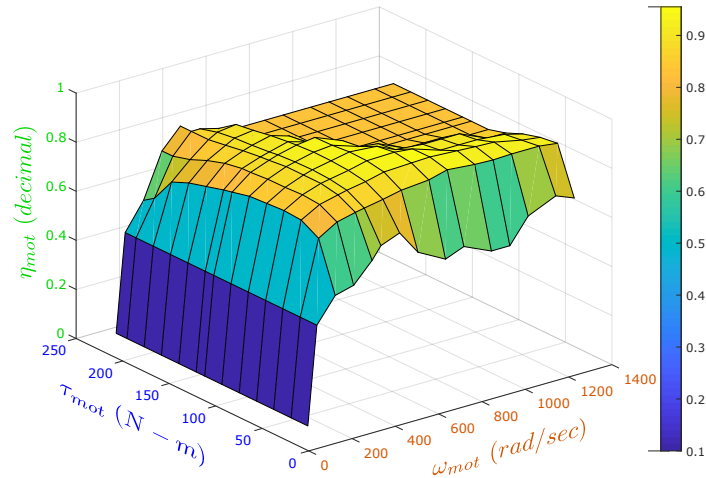


Figure 6.10: Tentative efficiency map of the traction motor used in Toyota Prius’s electrified powertrain.

corresponding to any row are carrying infinite immediate cost, it implies no feasible action available for that state.

### Simultaneous Q-value Iteration and Policy Iteration

A random Q-matrix is created with the same dimensions as the immediate matrix. The elements of the Q-matrix represent the Q-value of feasible state-action combinations. The Q-values are updated during sampling through collected drive cycles and generated drive cycles with the Q-learning algorithm. This subsection briefly corroborates how the matrix with random Q-values transforms into a matrix with near-global optimal Q-values. The Q-learning algorithm approaches the near-global optimal Q-value of a specific state-action combination as the RL agent interacts more and more with that state-action combination during the drive cycle. Hence, the RL agent is sampled through the collected and generated drive cycles multiple times so that each state-action combination can have multiple interactions with the RL algorithm. Ideally, 10000-50000 interactions are required for perfect convergence to global

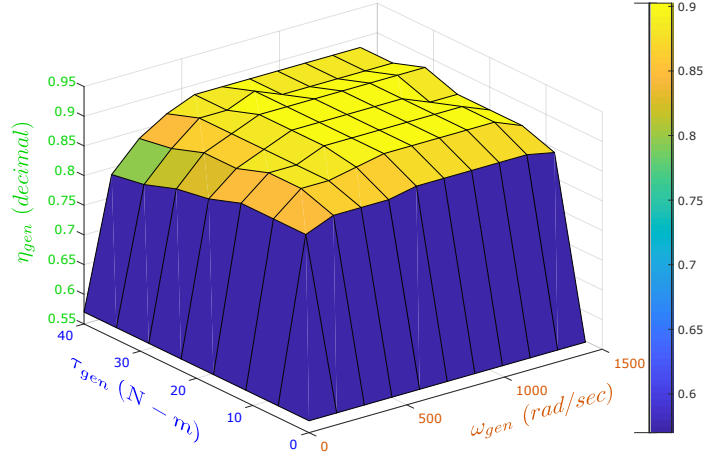


Figure 6.11: Tentative efficiency map of the generator used in Toyota Prius's electrified powertrain.

optimal Q-value for any state-action combination.

Since the goal is to have a near-global optimal control policy, Q-learning should be competent to culminate at the near-global optimal policy through the policy iteration technique. Q-learning technique leverages its intrinsic property of simultaneous Q-value iteration and policy iteration to achieve near-global optimal policy. The algorithm used for simultaneous iteration is shown in Fig.6.13,6.14. There are two key aspects of this implementation of Q-learning. First, the RL agent chooses a feasible action at any state by performing the following equation

$$a^*(t) = \arg \min_{a(t)} \left[ Q_k(s_t, a_t) \right] \quad (6.3.1)$$

with a probability of 95% and chooses a feasible action arbitrarily with a probability of 5%. This balances between exploitation-exploration inside of the policy iteration process. Second, the responsibility of completing every drive cycle with charge sustaining criteria is solely on the RL agent. If the battery depletes below 10% or recharged

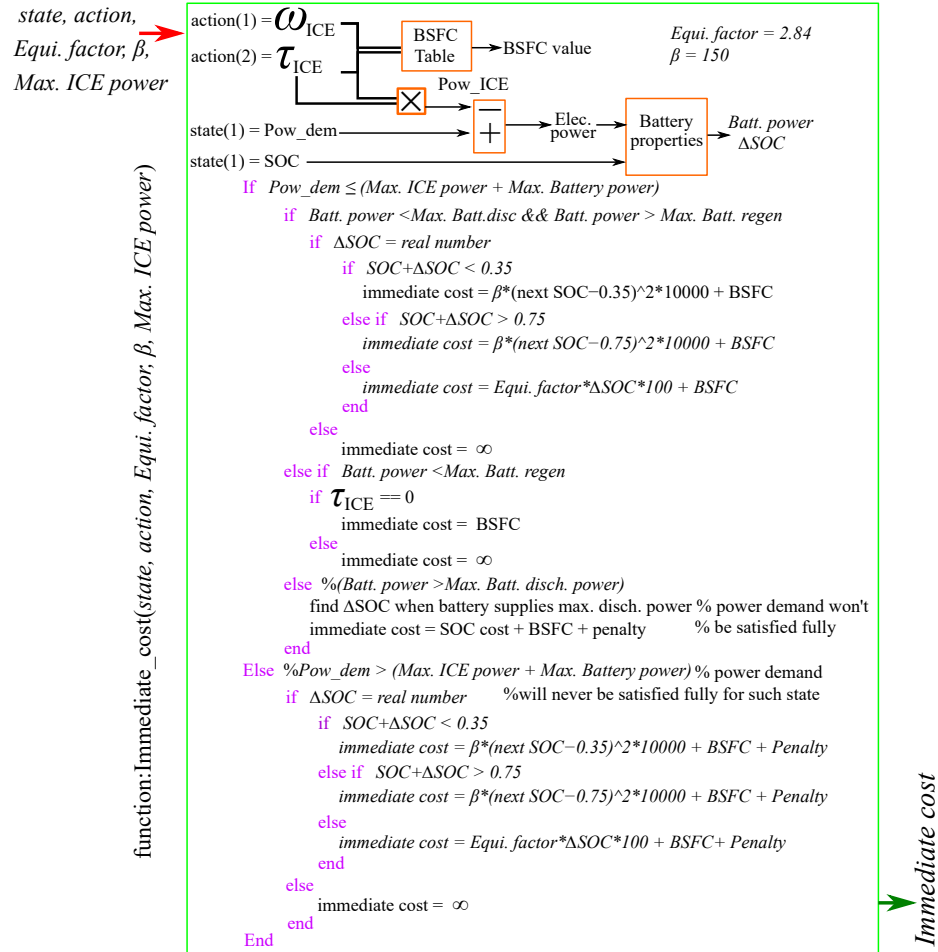


Figure 6.12: Algorithm used to calculate immediate cost for every combination of state and action.

beyond 90%, the RL will be considered unsuccessful in completing the drive cycle.

## 6.4 Performance of Energy Management System with Q-Learning based RL Agent

The Q-learning-based RL agent's first objective is to complete every drive cycle based solely on the RL algorithm, and the second objective will be to complete every drive



Table 6.2: Discretized values of Vehicle’s Power Demand, Battery SOC, ICE Torque, and ICE Speed

$Pow_1$	$Pow_2$	$Pow_3$	...	$Pow_{33}$	$Pow_{34}$	...	$Pow_{69}$	$Pow_{70}$	...	$Pow_{106}$
-52020	-30470	-25040	...	-1.667	300	...	18400	19211	...	123460
$SOC_1$	$SOC_2$	$SOC_3$	...	$SOC_{15}$	$SOC_{16}$	...	$SOC_{28}$	$SOC_{29}$	...	$SOC_{41}$
0.10	0.12	0.14	...	0.38	0.40	...	0.64	0.66	...	0.90
$State_1$	$State_2$	$State_3$	...	$St_{1327}$	$St_{1328}$	...	$St_{2816}$	$St_{2817}$	...	$state_{4346}$
-52020, 0.10	-52020, 0.12	-52020, 0.14	...	-1.667, 0.38	-1.667, 0.40	...	18400, 0.64	18400, 0.66	...	123460, 0.90
$ICE_1^\tau$	$ICE_2^\tau$	$ICE_3^\tau$	...	$ICE_{15}^\tau$	$ICE_{16}^\tau$	...	$ICE_{25}^\tau$	$ICE_{26}^\tau$	...	$ICE_{30}^\tau$
0	5	10	...	70	75	...	120	125	...	142
$ICE_1^\omega$	$ICE_2^\omega$	$ICE_3^\omega$	...	$ICE_{30}^\omega$	$ICE_{31}^\omega$	...	$ICE_{65}^\omega$	$ICE_{66}^\omega$	...	$ICE_{92}^\omega$
0	1000	1050	...	2400	2450	...	4150	4200	...	5500
$Action_1$	$Action_2$	$Action_3$	...	$Ac_{1318}$	$Ac_{1319}$	...	$Ac_{2273}$	$Ac_{2274}$	...	$Ac_{2472}$
0, 0	5, 1000	5, 1050	...	70, 2400	70, 2450	...	120, 4150	120, 4200	...	142, 5500

cycle with charge sustenance. Both of these primary objectives show a heavy reliance on the choice of “Immediate cost” corresponding to every feasible combination of state and action. As shown in Fig.6.15, cycle#1 and cycle#2 show improvement proportional to increase in training cycle. In both cycles, the RL-based agent can complete the full cycle after approximately 500 training iterations. However, the RL cannot always complete the full drive cycle even after 500 training iterations. This nature can be justified with the  $\epsilon - greedy$  control selection policy.

Moreover, for the perfect convergence of Q-values, more and more training iteration is required. The trend of final SOC proves that the Q-learning-based RL agent can maintain charge sustenance, and it reiterates the heavy reliance on the choice of “Immediate cost”. 35% SOC is intended as the final SOC in the “Immediate cost” formulation. The Q-learning-based RL agent can converge within the range of 40% to 35% after 800 training iteration. Fig.6.16 depicts the SOC profiles obtained from 11

```

*knnsearch() is the function used for nearest-neighbor
*Power_dem=f(vel,accel,grade)

for i=1:max_Iteration
  for jj=1:1:number of drive cycles
    SOC_cs = SOC_initial; *cs = current state, *ns = next state
    Pow_dem_cs = obtained from 1st time-step of drive cycle data;
    for t=1:1:length of each of the drive cycles
      index_power = knnsearch(state_power_index',pow_dem_cs);
      index_SOC = knnsearch(state_SOC_index',SOC_cs);

      cs = (indx_SOC-1)*length(state_power_index)+indx_power;
      Finding most probable current state

      Using Battery properties, calculate battery resistance, OCV,
      , max. dischar power, and max. regen power@ current SOC

      Finding an action for the current state:
      policy iteration is inherent to this finding process

      temp = rand(1);
      if temp < 0.1 % implementing ε-greedy policy
        n_actions = find(q(cs,:) < 50000);
        action_cs = n_actions(randi([1 length(n_actions)],1,1));
      else
        n_actions = find(q(cs,:) == min(q(cs,:)));
        action_cs = n_actions(randi([1 length(n_actions)],1,1));
      end
      eng_pow = action{action_cs}(3);

    continued.....
  end
end

```

Figure 6.13: Part.1: RL algorithm for simultaneous policy iteration and Q-value iteration.

arbitrary simulations for one of the real-world drive cycles with the trained Q-learning-based RL agent. The ICE speed and ICE torque profile for the first simulation is furnished in Fig.6.17 and Fig.6.18. One of the best performing RL agent's simulation after 800<sup>th</sup> training iteration is juxtaposed with DP-based simulation for performance evaluation of the trained agent. The result of the comparative study is depicted in Fig.6.19.

```

ElecPow = state_power_index(indx_power)-eng_pow;
Using ElecPow, battery resistance, OCV, and other battery
properties, ΔSOC is calculated
SOC_ns = SOC_cs - ΔSOC;
Pow_dem_ns = obtained from (t+1) time of drive cycle data;
indx_power = knnsearch(state_power_index',pow_dem_ns);
indx_SOC = knnsearch(state_SOC_index',SOC_ns);

Finding most probable next state
ns = (indx_SOC-1)*length(state_power_index)+indx_power;

Finding an action for the next state: Q-value iteration and
policy iteration is inherent to this finding process
n_actions = find(q(ns,:) < 50000);
temp = rand(1);
if temp < 0.1 % implementing ε-greedy policy
    min_q = q(ns,n_actions(randi([1 length(n_actions)],1,1)));
else
    min_q = 5000;
    for j=1:length(n_actions)
        min_q = min(min_q,q(ns,n_actions(j)));
    end
end

% Update q-values as per Bellman's equation
q(cs,action_cs)=q(cs,action_cs)+
    +alpha*(immed_cost(cs,action_cs)+
    +gamma*min_q-q(cs,action_cs));
pow_dem_cs = pow_dem_ns; SOC_cs = SOC_ns;
end
end

```

Figure 6.14: Part.2: RL algorithm for simultaneous policy iteration and Q-value iteration.

## 6.5 summary

It is shown in this chapter that RL can be employed to achieve a near-global optimal decision making policy for an EMS in real-world driving scenarios. Q-learning implementation steps are streamlined, starting from state variables, action variables, articulation of immediate consequence function up to systematic update Q-values. The powertrain architecture does not affect the choice of state variables. Only two

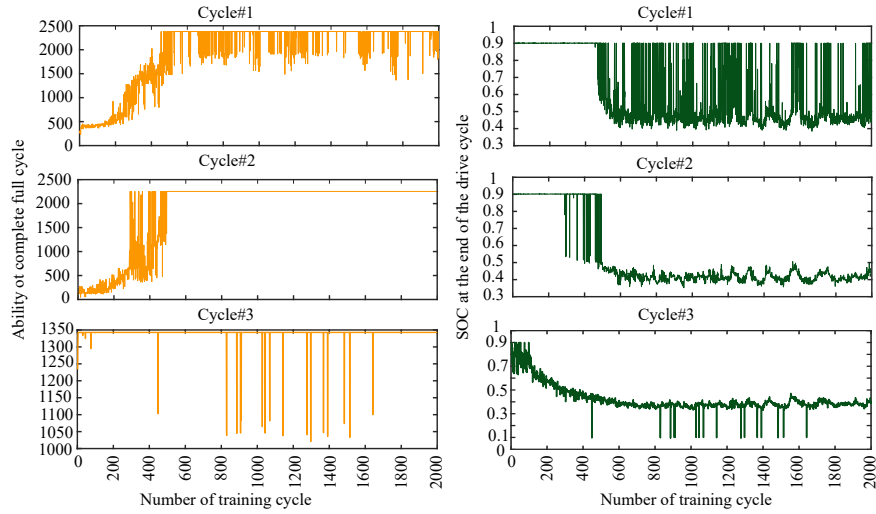


Figure 6.15: Result of Q-learning: Ability of full cycle completion with charge sustenance.

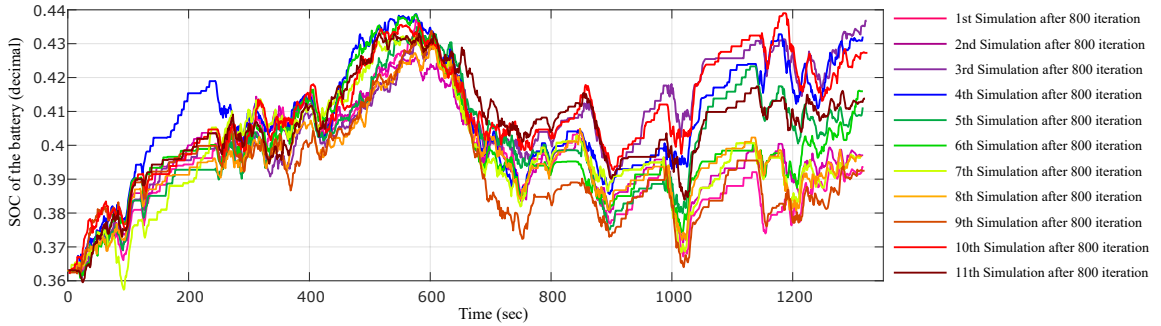


Figure 6.16: Trend of SOC profiles obtained from different simulations of a Q-learning-based RL agent trained with 800 iterations.

state variables are chosen for this chapter, and this number can be increased for a more accurate description of driving situations. However, the increased number of state variables increases computational complexity and retards the convergence rate of Q-learning. Such increase of computational complexity indicates the probable scope of a trade-off analysis between desired accuracy and computational complexity in the future. Since the powertrain architecture affects only the choice of action

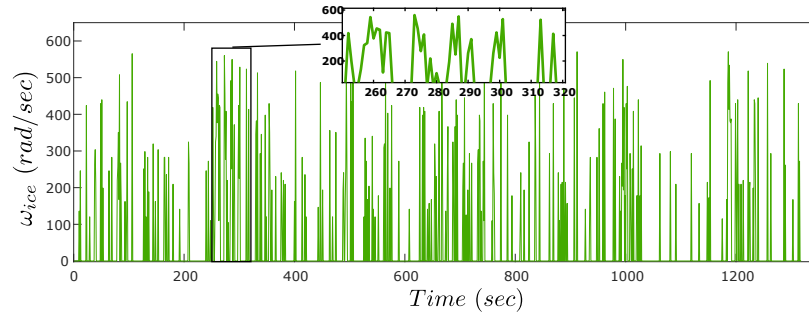


Figure 6.17: Time-series of engine speed obtained from a Q-learning-based RL agent trained with 800 iterations.

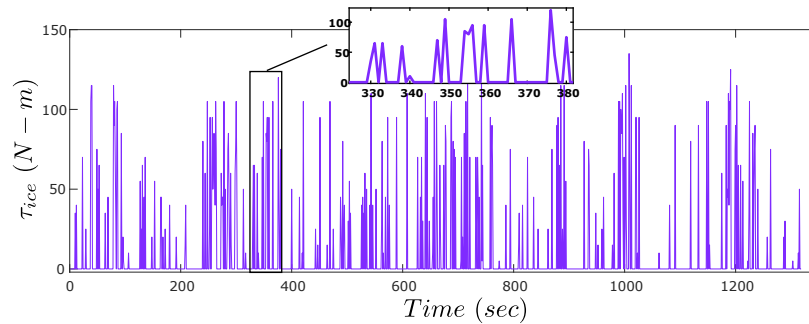


Figure 6.18: Time-series of engine torque obtained from a Q-learning-based RL agent trained with 800 iterations.

states, Q-learning can be pretty easily implemented to another powertrain architecture with minor modifications. The Q-learning has achieved 88% of the performance that would have been achieved by DP-based control policy.

Moreover, such performance is achieved through limited interaction and without pre-knowledge of the drive cycle. However, it is recommended to compare Q-learning and local optimal control policies to corroborate the superiority of the former over the latter. In conclusion, it can be said that Q-learning has a prospectus of becoming a near-global control policy for EMSs of the electrified powertrain in real-world driving scenarios.

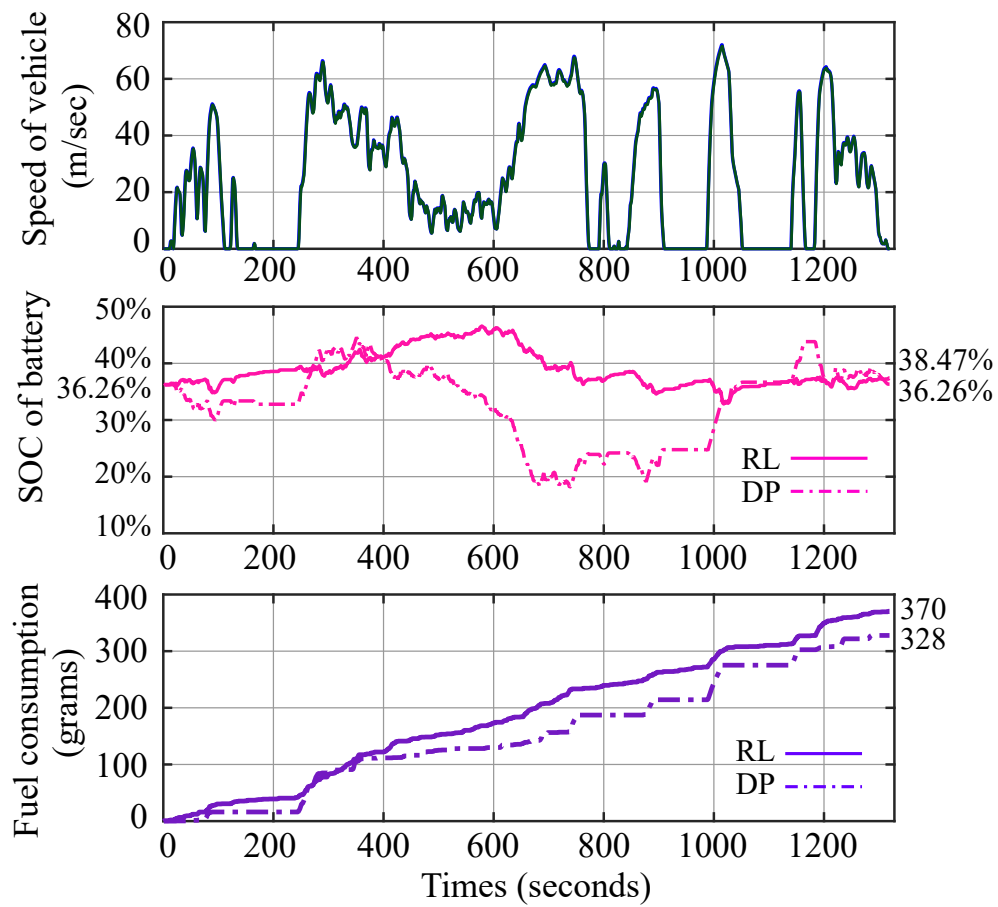


Figure 6.19: Comparison of performance between DP and a Q-learning-based RL agent trained with 800 iterations.

## **Chapter 7**

# **Application of Asynchronous Actor-Critic Agent in Real-time Energy Management of a Multi-mode Electrified Powertrain**

## 7.1 Introduction

The previous chapter established a sound conceptual foundation of RL making a smooth avenue for advanced reinforcement learning algorithm and their implementations in solving the energy management problem. A framework is presented in this chapter for implementing a state-of-the-art reinforcement learning agent, i.e., asynchronous advantage actor-critic (A3C) agent, as an energy management controller of a multi-mode electrified powertrain. The framework emulates the real-time interaction between the agent and the electrified powertrain and shows how efficient the agent can yield near-global optimal control for unfamiliar real-world driving scenarios. This chapter elucidates the asynchronous actor-critic agent’s effectiveness as a prospective real-time and real-world controller for a multi-mode electrified powertrain’s energy management system. This chapter posits the key novel contribution of this article. The real-time and real-world interaction between an energy management controller and continuously changing unfamiliar driving situations has been posing a critical challenge to the scholars who are working incessantly on finding an apposite control strategy yielding near-global optimal performance for unfamiliar drive cycles at the first interaction itself. The online simulation presented in this chapter corroborates that near-global optimal control is achievable in real-world driving scenarios if an asynchronous actor-critic agent is employed as the energy management controller in onboard vehicular control unit with a multi-core CPU.

A hybrid electric vehicle deployed with an intelligent energy management system can collect and store current driving data on-the-fly and generate “N” numbers of random drive cycles with the Markov chain model using the collected data. The agent can be trained with a simple multi-core CPU leveraging the agent’s asynchronous and



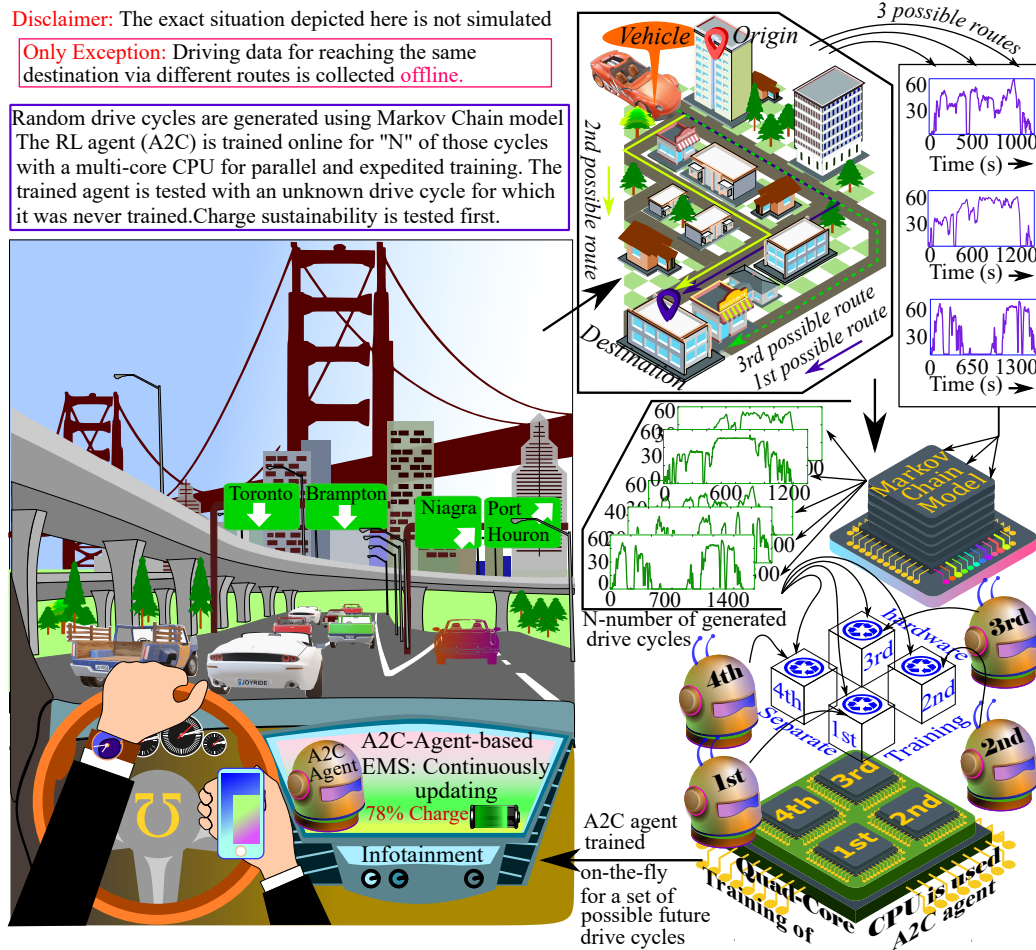


Figure 7.1: Main contribution of this dissertation: Application of asynchronous actor-critic in yielding near-global optimal control in real-world driving scenario.

parallel training capability. After training with 50 randomly generated drive cycles for 20 iterations, the agent is tested on another randomly generated drive cycle and a standard federal drive cycle to validate the agent's competency in unknown driving situations. The asynchronous and parallel training expedites the training process with quicker convergence. Although the data is not collected on-the-fly in the presented work, every other thing, including random generation of drive cycles from previously collected driving data, parallel training of the agent, and trained agent's testing, are

done online. The results obtained from the simulations corroborate the superiority of the agent over optimal equivalent consumption minimization strategy in terms of fuel-economy when applied on a random real-world drive cycle.

### 7.1.1 Highlights

- Solving energy management problem of a multi-mode hybrid electric vehicle.
- An asynchronous actor-critic agent is employed to solve energy management problem.
- Online training of the reinforcement learning agent with real-world drive cycles.
- The Reinforcement learning agent is tested on random drive cycles after training.
- Both training and testing are executed in model-in-the-loop simulation.
- Performance of the trained agent is compared with online optimal ECMS controls.

## 7.2 Inspiration for Online Simulation of Asynchronous Advantage Actor-Critic Agent

RL was first employed in solving the energy management problem of a parallel hybrid electric vehicle in [150]. An earlier article by the authors of this brief [151] and a handful of literature presented MATLAB®-based implementation of RL-based EMS

for an ePT [27, 152–160]. In [27], the authors presented the first use of RL in articulating the control strategy for the EMS with an ADVISOR-based vehicle simulation model. In [153], the authors have included an extra cost term in the immediate cost structure to procrastinate battery health degradation.

In [152], the authors have employed an ANN to replace the Q-table, and they have proposed the ANN as a wield to eliminate the “*curse of dimensionality*” that would have engendered from fine quantization in the Q-table approach. ANN helps eliminate the error from the approximation of continuous state and action with quantized states and actions [156]. The RL’s convergence rate improves, and the computational load reduces when an ANN is replaced with the DNN [157]. In [155], the authors have introduced the use of an ANN-based function approximator for articulating an actor-critic-based RL agent functioning as the EMS of an ePT. In both [158] and [159], the authors have architected the blueprint of real-time RL implementation using an online update of the TPM. In [159], the authors have proposed a framework for accelerated Q-learning in real-time with pre-calculated optimal control stored in look-up table (LUT)s. In [161], the authors have presented a sensitivity analysis on Q-learning parameters such as state and action discretization, state selection, and *exploration-exploitation* ratio. In [162], the authors have shown that using more state variables with coarse quantization is a more pragmatic way of describing the driving situations. Despite various irrefutable research contributions, all of these aforementioned literature lacks in one aspect, i.e., the online simulation or real-time emulation of the agent-environment interaction. Although the mathematical foundation remains the same for offline (MATLAB®-based) and online (Simulink®-based MIL) simulation, the implementation steps are not the same. The RL controller’s framework should

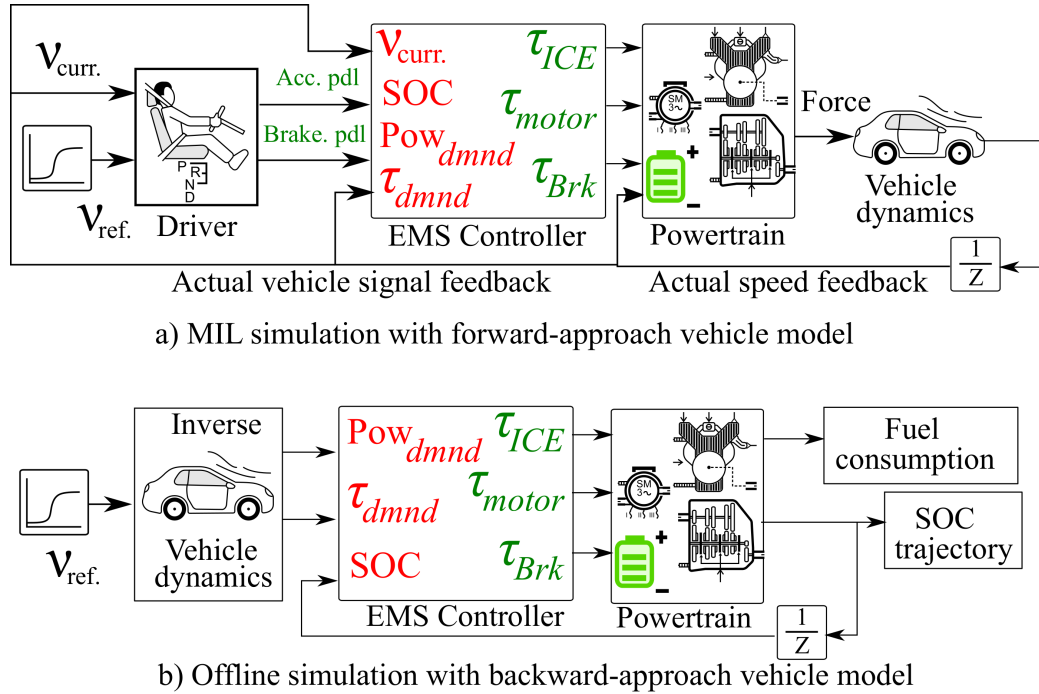


Figure 7.2: Schematic diagram of forward-approach and backward-approach simulation model.

be compatible with the MIL simulation platform’s real-time capable vehicle plant model. Needless to say that MATLAB <sup>®</sup>-based offline simulation cannot emulate the real-world interaction between a controller and the vehicle. In [163], the authors have presented a MIL simulation platform showing the interaction between the forward-approach simulation model (FSM) of a HEV and a Q-learning-based EMS for a pre-known drive cycle. Since Q-learning suffers from approximation error resulting from state and action variable’s discretization, it can handle either coarse discretization with more state variables or a fine discretization with fewer state variables. Undeniably, there is a requirement for a MIL simulation platform to emulate real-time interaction between an ANN-based RL controller and an FSM of an HEV.

### 7.3 Contribution

Looking at the necessity for an ANN-based application of RL agent as an online EMS for an ePT, the authors of this brief will present a DRL framework with an asynchronous actor-critic agent as an EMS controller for an ePT. Both the DRL agent and vehicle plant model are modeled in the Simulink <sup>®</sup> environment so that the real-time interaction between a DRL-based EMS and a physical HEV can be emulated.

Moreover, the A3C agent is applied for the first time in literature, to the best of the author's knowledge, for the energy management system of a multi-mode electrified powertrain. The DRL framework needs to be meticulously developed to consider additional feasibility constraints confronted by the multi-mode ePT.

The blueprint of yielding near-global optimal result in an unfamiliar driving situation through online training is testified if a simple multi-core CPU accompanies the onboard data collection system.

The DRL agent's performance will be juxtaposed to the performance obtained through a well accredited online control strategy, i.e., optimal ECMS. The DRL-based EMS will be accompanied by a low-frequency motor-torque-compensation control during all-electric to hybrid-electric mode-shift events and vice-versa.

## 7.4 Requirement and Fundamentals of Functional Approximation

Since the RL agent does not have prior knowledge of future states, it updates the policy from  $\pi$  to  $\pi'$  with an updated Q-value estimate corresponding to all state-action pairs. The prerequisite to such a policy-update process is quantifying the n-dimensional state-space and m-dimensional action-space into a  $\mathcal{N}$  and  $\mathcal{M}$  finite elements, respectively, and storage of all quantized state-action pairs' Q-value ( $Q : \mathcal{N} \times \mathcal{M} \mapsto \mathbb{R}$ ) in a repository, known as Q-table. However, the Q-table size grows exponentially, engendering “*curse of dimensionality*” as the discretization becomes more refined than before. The obvious and convenient way to circumvent such a problem is using a functional approximator, also known as compact parametric representation [164] ( $\tilde{Q}(s, a; \psi) \approx Q^*(s, a)$ ), where  $Q^*(.)$  is the optimal value function, and  $\psi$  is a parameter vector  $\psi \in \mathbb{R}^k$  [165]. Quintessentially, RL algorithm such as temporal-difference learning is used to continuously update the  $\psi$  vector's values to minimize the error between  $\tilde{Q}(.; \psi)$  and  $Q^*(.)$ .

### 7.4.1 Linear and Non-Linear Functional Approximation

One crucial factor in choosing the type of functional approximation is its differentiability, which is necessary for tracking its convergence and applicability of TD learning [25]. Linear combination of the feature vector and ANN are considered as well-known apposite functional approximator due to their conformity with the requirements mentioned above. Linear approximation of action-value and state-value functions can be expressed as a linear combination of a fixed set of feature functions

as follows [164]:

$$\begin{aligned}\tilde{Q}(s, a; \psi) &= \sum_{\lambda=1}^{\Lambda} \psi(\lambda) \Phi_{\lambda}(s, a) \\ &= \{\psi_1, \dots, \psi_{\Lambda}\}^T [\phi_1(s, a), \dots, \phi_{\Lambda}(s, a)]\end{aligned}\tag{7.4.1}$$

$$\begin{aligned}\tilde{V}(s; \psi) &= \sum_{\lambda=1}^{\Lambda} \psi(\lambda) \Phi_{\lambda}(s) \\ &= \{\psi_1, \psi_2, \dots, \psi_{\Lambda}\}^T [\phi_1(s), \dots, \phi_{\Lambda}(s)]\end{aligned}\tag{7.4.2}$$

$\psi \in \mathbb{R}^{\Lambda}$  and  $\Phi \in \mathbb{R}^{\Lambda}$  are parameter and feature-function vectors, respectively. Both of them have much less dimension than the gargantuan dimension of finely quantized state-space ( $\mathcal{S} \in \mathbb{R}^n$ ) or even the infinite dimension of continuous state-space ( $\mathcal{S}_{cont.} \in \mathbb{R}^{\infty}$ ). Hence, the linear approximation can be implied as a mapping from low-dimensional parameter space to high-dimensional state-space  $\mathbb{R}^{\Lambda} \mapsto \mathbb{R}^n$ . Feature vectors  $\Phi(s, a)$  and  $\Phi(s)$  contain most of the plausible functions representing salient features associated with an action-state combination and a state, respectively.

The feature functions of either  $\Phi(s, a)$  or  $\Phi(s)$  must be judiciously handpicked with expert knowledge, and that is the biggest challenge of using linear approximators for value functions in a complex decision-making process. Nonlinear approximator such as multi-layer perceptron (ANN) is a dexterous wield to circumvent the challenge mentioned above [164,166]. The architecture of a generalized nonlinear approximator is as follows:

$$\tilde{Q}(s, a; \psi) = g(\psi(\lambda) \Phi_{\lambda}(s, a))\tag{7.4.3}$$

, where  $g(\cdot)$  is a nonlinear function representing the architecture of the approximator. Craftiness must be attributed to the tailoring of multi-layer perceptron to stifle some of the malicious effects engendered due to the complexity of ANN and amplify the approximation accuracy [166]. Essentially, an ANN-based approximator’s performance is governed by a tactical compromise between network complexity and aspiration of approximation accuracy [167].

### 7.4.2 Learning of Approximator Parameters

The core concept of this entire function approximation based RL framework lies in the learning algorithm employed for updating the parameter vector  $\psi \in \mathbb{R}^\Lambda$ . Undeniably, TD learning would be the most appropriate algorithm to update the parameter vector. Unsupervised learning, which is also known as model-free learning or prediction-based learning, is the cornerstone for the future of AI. TD learning has proclaimed its ubiquitous presence in almost all unsupervised learning algorithms because of its most generalized structure. Supervised learning is applied when the targets of the function approximation are fixed. In contrast, unsupervised learning is indispensable when there is no explicit value of the approximator’s targets available or vacillating targets.

The predicted value function  $\tilde{V}(s; \psi)$  of any non-terminal state  $S_t$  is compared with the predicted value function of the subsequent state  $S_{t+1}$  instead of comparing it with the optimal value function  $V^*(S_t)$  of  $S_t$  due to lack of explicit knowledge about  $V^*(S_t)$ .

$$\delta_t = r_t + \gamma \tilde{V}(S_{t+1}; \psi) - \tilde{V}(S_t; \psi) \quad (7.4.4)$$

$\delta_t$  quantifies the *TD error* between two consecutive predictions. In simpler words,



TD learning is the process of updating a guess based on another guess. Leveraging this fact, TD learning does not need to wait until the outcome of the episodic sequential decision-making process to update any non-terminal state’s value function. The TD learning is nothing but an articulation of a rule for updating the parameter vector so that the approximator can predict every state’s value function with utmost accuracy. Quintessentially, the parameter vector of the approximator is updated with a *stochastic gradient – descent*(SGD) method [25]:

$$\begin{aligned}
 \psi_{t+1} &= \psi_t - \frac{1}{2}\varphi\nabla\left[r_t + \gamma\tilde{V}(S_{t+1};\psi_t) - \tilde{V}(S_t;\psi_t)\right]^2 \\
 &= \psi_t + \varphi\left[r_t + \gamma\tilde{V}(S_{t+1};\psi_t) - \tilde{V}(S_t;\psi_t)\right]\nabla\tilde{V}(S_t;\psi_t) \\
 &= \psi_t + \varphi\delta_t\nabla\tilde{V}(S_t;\psi_t)
 \end{aligned} \tag{7.4.5}$$

, where  $\varphi$  quantifies a learning rate and  $\nabla f(\psi)$  refers to a vector of partial derivatives with respect to the weight vector as follows:

$$\nabla f(\psi) \doteq \left(\frac{\partial f(\psi)}{\partial \psi_1}, \frac{\partial f(\psi)}{\partial \psi_2}, \dots, \frac{\partial f(\psi)}{\partial \psi_\Lambda}\right)^T \tag{7.4.6}$$

RL can be broadly categorized into three groups, i.e., actor-based, critic-based, and actor-critic-based algorithms [168]. Actor-based algorithms focus only on policy iteration to optimize the performance of the agent (controller). In contrast, critic-based algorithms focus explicitly on the iteration of value function to optimize the agent and implicitly focus on selecting action through either “*Greedy-search*” or “ $\epsilon$ -*Greedy-search*” method. Actor-critic-based algorithms are developed by amalgamating the advantages of both actor and critic. Actor-critic explicitly employs two separate

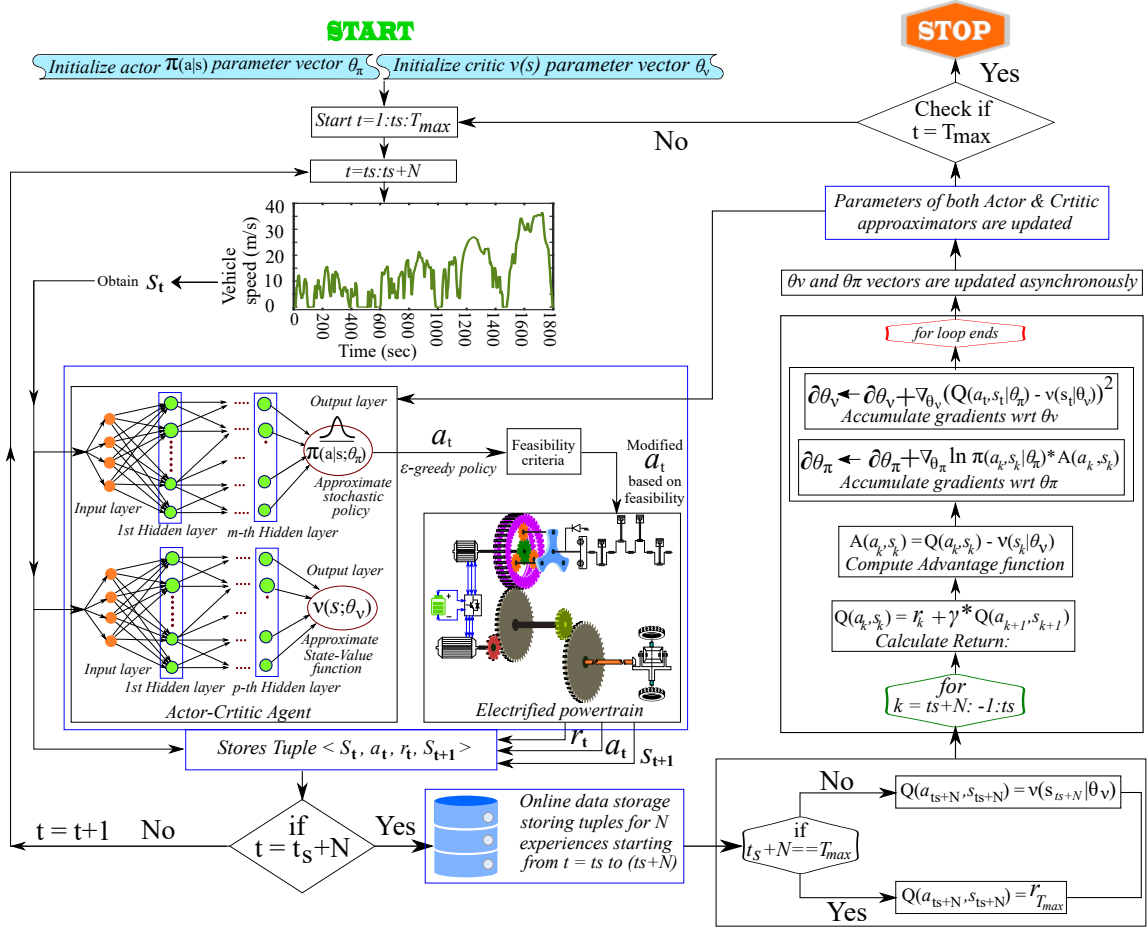


Figure 7.3: Schematic diagram of actor-critic based agent with nonlinear function approximation.

networks to optimize the policy function and value function separately. The Actor-network of the actor-critic agent updates the policy function, and the critic-network evaluates the goodness of the advocated policy [168].

## 7.5 Training of Deep Reinforcement Learning Agents

The success of function approximation highly relies on the states' sampling process and on the process of updating the parameters [25]. In [165], the authors have reiterated that the states should be sampled strictly with steady-state probabilities of Markovian dynamics during the functional approximator's training epochs. In the absence of steady-state probabilities, the only remaining option is to train the approximator with online sampling [25].

In [165], the authors not only have expressed their skepticism on the stability of the well-accredited RL algorithm, i.e., TD(0), when used for updating the parameters of an ANN-based nonlinear functional approximator but even insinuated toward conceiving divergence in both value-function approximations and parameters of the approximator. The prime reasons for such instability and divergence can be enumerated as follows:

- The sampled states in a Markov decision process are highly correlated to each other, i.e., a transition between a present state and its immediate predecessor state cannot be modeled separately. The transition dynamics between two consecutive states are highly correlated with all the previous transitions [26,169].
- Quintessentially, an update of the parameter vector focused on reducing the difference between optimal and approximated value function corresponding to a specific region of state-space can have an adverse effect on gap reduction corresponding to other state-space regions [170].

### 7.5.1 Experience Replay

A promising solution to the instability mentioned above and divergence had already been prescribed long before in 1992 by the authors of [171]. They advocated utilizing the past experiences judiciously to update the parameter vector of functional approximation. The underlying concept of *experience replay*, as introduced in [171], is buttressed with two phenomena, i.e., the storage of past experiences followed by presenting those stored experiences to the agent at every occasion whenever it would like to perform the value iteration through a learning algorithm. An *experience* is defined as a quadruple  $(\langle S_t, A_t, r_t, A_{t+1} \rangle)$ , i.e., state, action, immediate consequence, and the next state. In [170], the authors have proposed a method called *Neural Fitted Q Iteration*, leveraging the concept of *experience replay* for updating the parameter vector of a multi-layer perceptron. There is a clear guideline, elucidated in [172], for deciding which part of the agent’s past experiences should be catered to during the current update. It is elucidated that the RL agent should be catered with only the off-policy past experiences.

Another crucial aspect, known as *mini-batches*, which might not be indispensable for reducing the malicious effect of highly correlated states on stability, but still plays a cardinal role in expediting the convergence of an ANN-based functional approximator. In [26, 169], the authors have introduced the method of sampling *mini-batches* randomly from the stored *experience replay* instead of using the entire *experience replay* while updating the parameter vector.

### 7.5.2 Asynchronous Training of Advantage Actor-Critic Agent

Although *experience replay* has significantly solved the divergence problem and also circumvented the issue of correlated states, its implementation instigates another constraint, i.e., only off-policy training algorithms are suitable while using *experience replay*. On-policy sampling cannot be used to update the parameter vector. Furthermore, standard multi-core processing units might not suffice the memory requirement to facilitate *experience replay*, and hence advanced hardware like graphical processing unit (GPU) should be employed to perform the training of RL agent. In [173], the authors have brought a paradigm shift in reinforcement learning by introducing a disruptive concept called asynchronous training. Two cardinal virtues of asynchronous training, as elucidated in [173], are as follows:

- Asynchronous and simultaneous training of agents with multiple replicates of the environment, in which different parts of the state-space can mitigate the adverse effect of highly-correlated states.
- Actors can be equipped with distinct exploration policies, which can establish complete stabilization in the sampled states.

With these two properties mentioned above, the asynchronous method can enable an agent's on-policy learning, eliminating the required storage space for *experience replay*. Hence, specialized hardware is no longer required. Instead, a multi-core CPU can facilitate asynchronous training. Given these optimized properties, the A3C agent was applied as a control strategy for the EMS of an ePT for real-time emulation. Before delving further into the working principle of the A3C agent, it might be useful to consolidate all stages of this evolution from an ANN-based RL-agent to an A3C

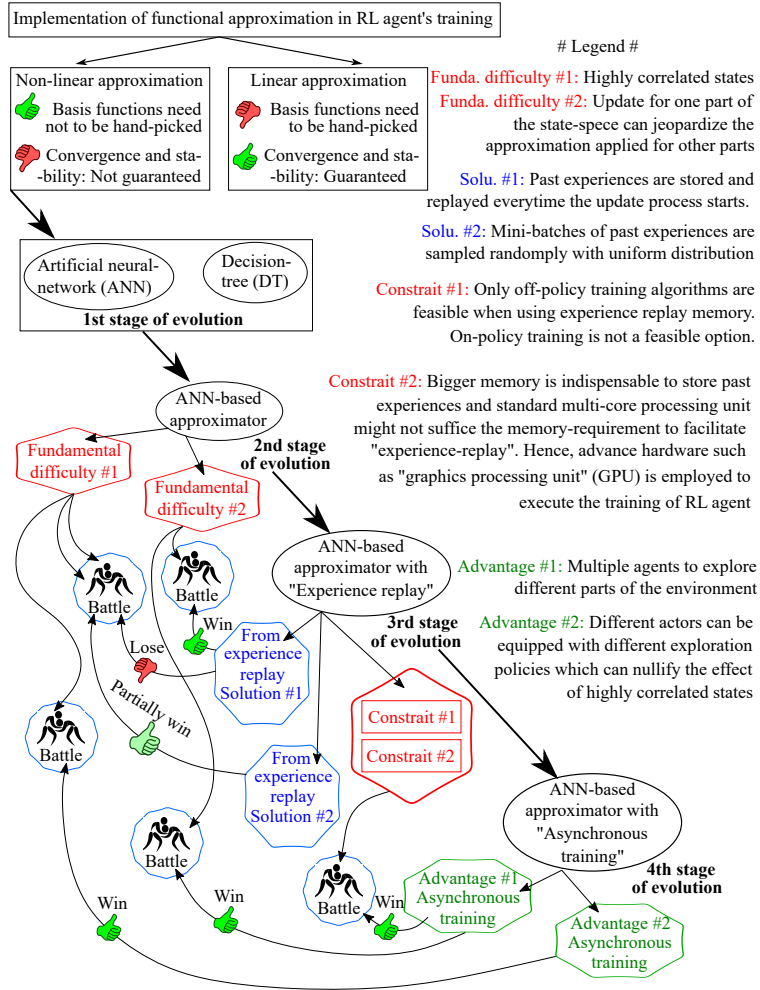


Figure 7.4: Evolution of ANN-based function approximator as RL agent.

agent. The process is depicted in Fig.7.4. As depicted in Fig.7.3, the asynchronous actor-critic algorithm starts with a random parameter vector for both actor ( $\psi_\pi$ ) and critic ( $\psi_\nu$ ). Since there are multiple replicates of the environment, it is usual that the algorithm shown in Fig.7.3 is applicable for all replicates of the environment. According to article [173], asynchronous actor-critic samples “ $N$ ” number of *state – action – next state – reward* interactions before each update of  $\psi_\pi$  and  $\psi_\nu$ . Suppose, the sampling starts at  $t_s^{th}$  time-step and continues up to  $(t_s + N)^{th}$  time-step. At

each time-step within this sampling period, the actor-critic agent receives the state information ( $s_t$ ) from the driving mission. The critic evaluates the current state's value function, and the actor decides an apposite action ( $a_t$ ) for the current state based on a  $\epsilon$ -greedy policy. In this aspect, it is worth mentioning that each replicate of the environment is equipped with a distinct  $\epsilon$ -greedy policy. As an immediate consequence of the action on the environment, i.e., the electrified powertrain, the agent receives a reward ( $r_t$ ), and the environment makes a transition to a different state ( $s_{t+1}$ ) at the next time-step. Both  $r_t$  and  $s_{t+1}$  are perceived by the agent in the next time-step. The tuple comprised of “state – action – next state – reward” is stored temporarily for the period mentioned above.

The estimated return after  $(t_s + N)^{th}$  time-step is zero if the  $(t_s + N)^{th}$  time-step is the end of the episode. The estimated return after  $(t_s + N)^{th}$  time-step is expressed as  $Q(s_{t_s+N}, a_{t_s+N})$  if that is a non-terminal time-step. Now, it is time to update the parameter vectors  $\psi_\pi$  and  $\psi_\nu$ . Before the “ $N$ ” steps, estimated return from state  $s_{t_s}$  ( $\mathbb{E}(\mathcal{R}_{t_s})$ ) was  $V(s_{t_s}; \psi_\nu)$ . But, after “ $N$ ” steps of sampling,  $\mathbb{E}(\mathcal{R}_{t_s})$  can be re-written in the following way:

$$\begin{aligned} \mathbb{E}(\mathcal{R}_{t_s}) &= r_{t_s} + \gamma r_{t_s+1} + \dots + \gamma^{N-1} r_{t_s+N-1} + V(s_{t_s+N}; \psi_\nu) \\ &= \sum_{k=1}^N \gamma^{k-1} r_{t_s+k-1} + V(s_{t_s+N}; \psi_\nu) \end{aligned} \tag{7.5.1}$$

Similarly,  $\mathbb{E}(\mathcal{R}_{t_s+1})$  for state  $s_{t_s+1}$  can be expressed before and after “ $N$ ” samplings

with Eq.7.5.2a and Eq.7.5.2b, respectively.

$$\mathbb{E}(\mathcal{R}_{t_s+1})_{before} = V(s_{t_s+1}; \psi_\nu) \quad (7.5.2a)$$

$$\mathbb{E}(\mathcal{R}_{t_s+1})_{after} = \sum_{k=2}^N \gamma^{k-2} r_{t_s+k-1} + V(s_{t_s+N}; \psi_\nu) \quad (7.5.2b)$$

Similarly,  $\mathbb{E}(\mathcal{R}_{t_s+2})$  for state  $s_{t_s+2}$  can be expressed before and after “ $N$ ” samplings with Eq.7.5.3a and Eq.7.5.3b, respectively.

$$\mathbb{E}(\mathcal{R}_{t_s+2})_{before} = V(s_{t_s+2}; \psi_\nu) \quad (7.5.3a)$$

$$\mathbb{E}(\mathcal{R}_{t_s+2})_{after} = \sum_{k=3}^N \gamma^{k-3} r_{t_s+k-1} + V(s_{t_s+N}; \psi_\nu) \quad (7.5.3b)$$

And similarly,  $\mathbb{E}(\mathcal{R}_{t_s+N-1})$  for state  $s_{t_s+N-1}$  can be expressed before and after “ $N$ ” samplings with Eq.7.5.4a and Eq.7.5.4b, respectively.

$$\mathbb{E}(\mathcal{R}_{t_s+N-1})_{before} = V(s_{t_s+N-1}; \psi_\nu) \quad (7.5.4a)$$

$$\mathbb{E}(\mathcal{R}_{t_s+N-1})_{after} = r_{t_s+N-1} + V(s_{t_s+N}; \psi_\nu) \quad (7.5.4b)$$

The intent of expressing Eq.7.5.1 to Eq.7.5.4b is to show that the value functions corresponding to  $s_{t_s}$ ,  $s_{t_s+1}$ ,  $s_{t_s+2}$ , ...,  $s_{t_s+N-2}$ , and  $s_{t_s+N-1}$  get an  $N$ -step,  $(N-1)$ -step,  $(N-2)$ -step, ..., two-step, and one-step update, respectively after a “ $N$ ”-step sampling period.

The gradient of  $\psi_\pi$  is calculated by utilizing the estimated advantage of applying action  $a_t$  at the state  $s_t$  [174]. The *advantage* is computed by subtracting the value function of state  $s_t$  ( $V(s_t; \psi_\nu)$ ) from the expected return ( $\mathbb{E}(\mathcal{R}_t)$ ) obtained through



applying an action  $a_t$  at the state  $s_t$ , as shown in Eq.7.5.5.

$$\begin{aligned} A(a_t, s_t; \psi_\pi) &= Q(a_t, s_t) - V(s_t; \psi_\nu) \\ &\approx \mathbb{E}(\mathcal{R}_t) - V(s_t; \psi_\nu) \end{aligned} \quad (7.5.5)$$

Now, the gradient of *advantage* function, as given by Eq.7.5.6

$$\nabla(\textit{advantage}) = (\nabla_{\psi_\pi} \ln \pi(a_t|s_t; \psi_\pi) A(a_t, s_t; \psi_\pi)) \quad (7.5.6)$$

, is employed to update the parameter vector of actor-network ( $\psi_\pi$ ). The gradient of  $(\mathbb{E}(\mathcal{R}_t) - V(s_t; \psi_\nu))^2$  is used to update the parameter vector of critic-network  $\psi_\nu$ . The gradients obtained from all time-steps within an N-step sampling period are garnered and summed to perform one single update on both  $\psi_\pi$  and  $\psi_\nu$ , as shown in Eq. (7.5.7a) and (7.5.7b).

$$\partial \hat{\psi}_\pi = \sum_{k=1}^N (\nabla_{\psi_\pi} \ln \pi(a_k|s_k; \psi_\pi) A(a_k, s_k; \psi_\pi)) \quad (7.5.7a)$$

$$\partial \hat{\psi}_\nu = \sum_{k=1}^N \frac{\partial (\mathbb{E}(\mathcal{R}_k) - V(s_k; \psi_\nu))^2}{\partial \psi_\nu} \quad (7.5.7b)$$

## 7.6 Multi-mode Powertrain Modeling for RL Application

A3C's application in solving the energy management problem for an ePT has been witnessed sparsely across the literature. While a handful of authors have presented RL's application in an electrified powertrain with parallel architecture [150, 152, 153,

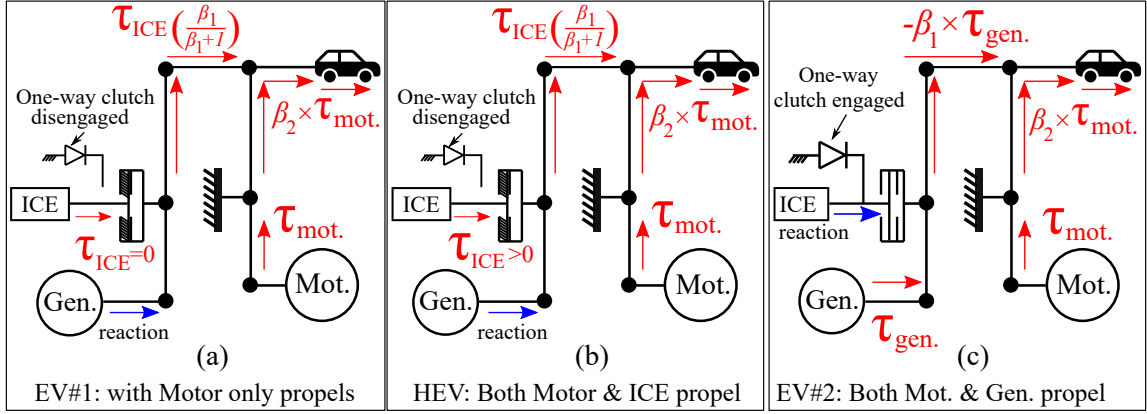


Figure 7.5: Schematic diagram of three distinct modes facilitated by the eVT transaxle.

156], a few authors have proffered its application in series architecture [159, 163] and power-split hybrid architecture [157, 175]. However, to the best of the authors' knowledge, there is no published work where RL is applied to a multi-mode ePT. The subsequent sections will present the justification behind taking a different approach for implementing RL for a multi-mode ePT.

The midsize 2575 kg representative passenger vehicle with a multi-mode e-VT [42] is selected for this study. The powertrain's cardinal components with multi-mode e-VT are an ICE, two EMGs, two PG-sets, and an HVB pack, as shown in Fig.7.6.

The only difference between the ePT shown in Fig.7.6 and the powertrain shown in chapter 6 is the presence of an extra all-electric mode with both traction motor and generator propelling the vehicle. This multi-mode electrified powertrain is also used in chapter 4 and chapter 5 for articulating local optimization-based EMS control.

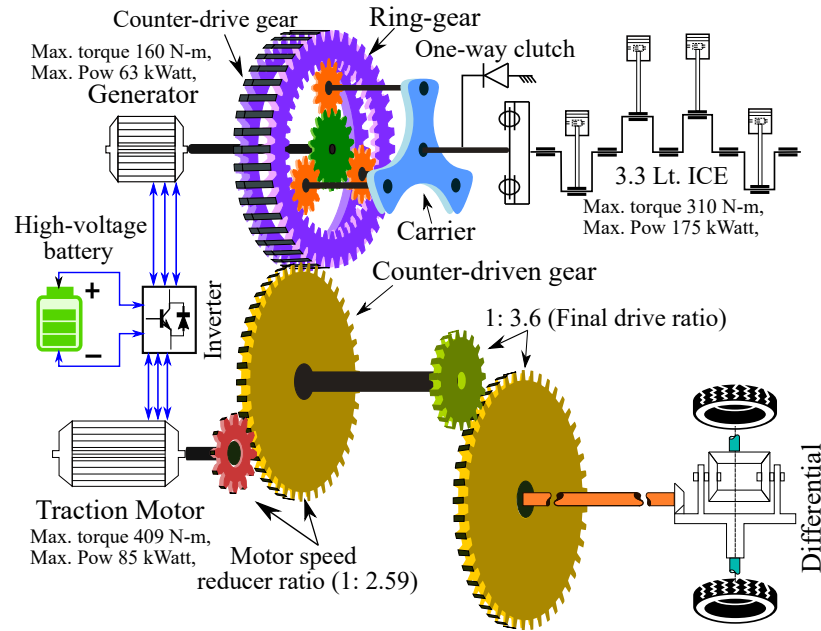


Figure 7.6: Schematic diagram of multi-mode eVT transaxle.

### 7.6.1 Motor Torque Compensation control during Mode-shifts

The generator starts applying positive torque to crank the ICE as soon as the EMS controller commands for all-electric to hybrid-electric mode-shift. All-electric to hybrid-electric mode-shift is highly susceptible to engine ripple torque due to low-speed operation of ICE. In [176], the authors have prescribed a simplified versions of dynamic coordinated control strategy (DCCS) to cope up with the adverse effect of engine ripple torque on transmission output, predominantly during all-electric to hybrid-electric mode-shift in power-split architecture. In this article, the high-frequency engine ripple torque is replaced by low-frequency mean-value engine resistance torque, as shown in Fig.4.11.c.

During cranking, the generator takes the responsibility of providing enough positive torque to crank the ICE, and traction motor takes the rest of the responsibility

in terms of sufficing torque demand at the transmission output. The expressions of coordinated torque for generator and traction motor during all-electric modes to hybrid-electric mode are given in Eq.7.6.1 and Eq.7.6.2, respectively.

$$\tau_{gen|coordinated} = \frac{\left\{ J_{iceeq} (\dot{\theta}_{icerqst} - \dot{\theta}_{iceplant}) - \tau_{fricice} \right\}}{(\beta_1 + 1)} + J_{geneq} \ddot{\theta}_{gen} \quad (7.6.1)$$

$$\tau_{mot|coordinated} = \frac{\beta_1 \left\{ \tau_{fricice} - J_{iceeq} (\dot{\theta}_{icerqst} - \dot{\theta}_{iceplant}) \right\}}{\beta_2(\beta_1 + 1)} + J_{moteq} \ddot{\theta}_{mot} + \frac{\tau_{out}}{\beta_2} \quad (7.6.2)$$

The engine stopping phase during hybrid-electric to all-electric mode-shift is equally victimized to the effect of engine resistance torque due to the ICE's operation below idle speed [177]. The expressions for compensation torque of generator and traction motor remain the same as Eq.7.6.1 and Eq.7.6.2. Next chapter will delineate the detailed application of coordinated control required during mode-shift events in a multi-mode ePT.

## 7.7 Online Implementation of Asynchronous Advantage Actor-Critic Agent as EMS Controller

There are multiple steps associated with the online implementation of an asynchronous actor-critic agent as a near-optimal EMS controller. All the steps are delineated as follows:

### 7.7.1 Forward Simulation Model of Vehicle Plant

RL is a *model-free* algorithm, and it is noteworthy to mention that the *model* does not refer to the vehicle plant model. Instead, it refers to the model of Markov decision problem (MDP). The model of an MDP is constituted with two elements, i.e., the probability of transitioning from one particular state to another state (state transition probability matrix) and an immediate consequence function (either reward or cost). Immediate consequence function is indispensable to both *model-based* and *model-free* RL algorithms. *Model-free* RL algorithms are the only wield for solving MDPs if the state transition probability matrix is not available. Needless to say that the state transition probability matrix cannot be computed unless the RL agent knows the entire drive cycle a priori. Therefore, *model-free* RL algorithms such as *Q-learning*, *SARSA*, and *Actor-Critic* become apposite when future driving situations are unknown before real-time driving. In a nutshell, a *model-free* RL algorithm in an HEV only needs to know the immediate consequences such as instantaneous fuel consumption, instantaneous tailpipe emission, and battery SOC from the real HEV to optimize its EMS strategy. Nevertheless, neither the immediate consequences mentioned above can be obtained from a real HEV in a MIL simulation, nor an RL-based algorithm will be used as an EMS strategy of a real HEV in the development phase of the RL agent. Therefore, a MIL simulation must have an FSM-based vehicle plant, as shown in Fig.7.2, for emulating real-time immediate consequence functions mentioned above.

### 7.7.2 Immediate Consequence Function

The articulation of the immediate consequence function plays a significant role in the convergence of the A3C agent. Since HEV is the application vehicle in this

brief, only fuel consumption and strict charge sustenance are focused at the end of the drive cycle. Only fuel consumption minimization is chosen as the EMS's prime objective to retain the simplicity of the A3C agent-based EMS at this nascent stage of development. Hence, the immediate consequence function for the A3C agent is as follows:

$$\begin{aligned} (\text{Immediate cost})|_{\Delta t} = C_1 * (MFR)_{fuel} * \Delta t + \\ C_2 * (SOC_{ref} - SOC_{curr.})^2 \end{aligned} \quad (7.7.1)$$

There is no thumb rule available for deducing these two constants ( $C_1$  and  $C_2$ ) since their value depends on the overall simulation model. Their values are quite important in achieving convergence of the A3C agent. For this analysis, their values are obtained by the trial-and-error method.  $C_1$  is considered to be zero, and  $C_2$  is considered to be one in this analysis. Both the objectives, i.e., fuel consumption minimization and charge sustainability, can be satisfied with just the second term in the cost function. Although the global optimality of the solution is not guaranteed even if it achieves charge sustainability, the solution will be very near to the global optimal as per Pontryagin's minimization problem if it can achieve charge sustenance. There is a scope of applying any optimization technique such as GA to find out optimum values for these two constants.

### 7.7.3 Implementation of A3C Agent in Simulink®

This article's center of attraction is the emulation of online interaction between A3C agent and the vehicle plant model in Simulink®. Similar to the vehicle plant model, the A3C agent is developed solely in the Simulink environment®.

### **Selection of State and Action Variables**

The first step of the agent’s design procedure is the selection of state and action variables. An A3C agent can receive only a continuous state variable from the environment but can emit an action variable in either continuous or discrete space. In this article, five state variables, i.e., SOC of the HVB, vehicle’s speed, vehicle’s acceleration, drive’s requested power, and road grade information, have been used. Tab.7.1 will enumerate every state variable’s lower and upper boundaries. The three state variables except the road grade and battery SOC can be extracted from the vehicle’s speed trajectory. In the current endeavor, only these five states are employed to represent the vehicle’s driving state. However, the driving state can be more distinctly identified if more state variables are used. More state variables such as ambient temperature, battery temperature, wind speed, traffic information, approximate remaining distance from the destination can be included in the state vector as per the required fidelity and objectives of the EMS. Vehicle’s speed and battery SOC provide reasonable accurate information about the driving to an EMS with the objective just fuel consumption minimization and charge sustainability of the battery. State variables such as ambient temperature and battery temperature have been proven to be useful information for an EMS with the additional objective of battery life prolongation. The increased number of state variables does not complicate the optimization of the DNN, since DNN’s convergence is not deteriorated by the additional state variables. However, this statement is only valid if the architecture of the DNN is retained same as previous.

As far as the selection of action variables are concerned, here it will be elucidated why articulation of an RL agent-based EMS is unique for the multi-mode ePT. As

mentioned earlier, the multi-mode e-VT selected for this article facilitates three distinct modes. Any given EMS control strategy for multi-mode ePT should decide the electrified mode followed by the control action(s) based on the DOF of the decided mode. As mentioned earlier, single-motor EV and hybrid-electric modes share the same kinematic relationships among all the prime-movers, and two-motor EV mode offers an entirely different kinematic relationship. Single-motor EV does not offer any DOF, and hence, the drive's torque demand must be satisfied by traction motor only. The ICE is not activated, and hence, both speed and torque of ICE are zero in the single-motor EV mode. Hybrid-electric mode offers two DOFs in the form of ICE speed and ICE torque. Two-motor EV mode offers one DOF in the form of traction motor's torque. Tab.7.2 enumerates the feasible combinations of control variables in different modes and their constraints.

The RL agent randomly selects the action variables, and the random combination of the action variables must be kinematically feasible for the randomly selected mode. There are four control variables considering all the three modes, but all four actions cannot be simultaneously applied to either of them. Traditionally, two layers of control are employed in an EMS to tackle this kind of situation with a multi-mode ePT, i.e., the outer layer of the control strategy selects the mode with some premeditated rule, and the inner layer optimizes either power-split or torque-split based on the DOF of the selected mode. Two separate RL agents can be employed to achieve two layers of control. However, it will be difficult to achieve convergence with two RL agents, given the fact that even a single RL agent can pose an ample amount of difficulty in tuning its hyper-parameters, which will be discussed shortly.



All the hindrances mentioned above arise for continuous action-space. These hindrances can be circumvented using discrete action-space instead of continuous, but with an extra amount of manual work, i.e., all feasible combinations of four control variables must be pre-computed and fed to the A3C agent. The feasibility criteria, elaborated in Tab.7.2, elucidate the rules for screening out feasible combinations of four action variables.

Tab.7.3 enumerates a handful of random samples of feasible and infeasible combinations of action variables for specific powertrain parameters such as  $\beta_1 = 2.87$  and  $\beta_2 = 2.5$ , and particular output torque  $\tau_{out} = 150N - m$ . Although this approach seems enticing, it is laden with procrastinated convergence due to the gargantuan size of the feasible action-space. For example, the number of feasible combinations of discrete actions can reach approximately up to 2700, even with a coarse discretization of  $\tau_{ice}$ ,  $\tau_{mot}$ ,  $\omega_{ice}$ , and *Mode*. It becomes a strenuous job for the RL agent to obtain even decent performance with such a gargantuan action-space.

### **Implementation of RL featured by ECMS**

Now, it is the perfect time to prescribe an impressive idea to conflate the concept of ECMS [15] for multi-mode ePT into the online implementation of an A3C agent. Observing the A3C agent’s ineptness alone in tackling 4-D action-space in both continuous and discrete format, especially for multi-mode ePT, it is proposed to buttress the A3C agent with an ECMS controller, which is competent enough to deal with a multi-mode ePT. Algo.1 furnishes the pseudo-code for instantaneous optimization-based ECMS for a multi-mode ePT. ECMS suits quite well for multi-mode EMS

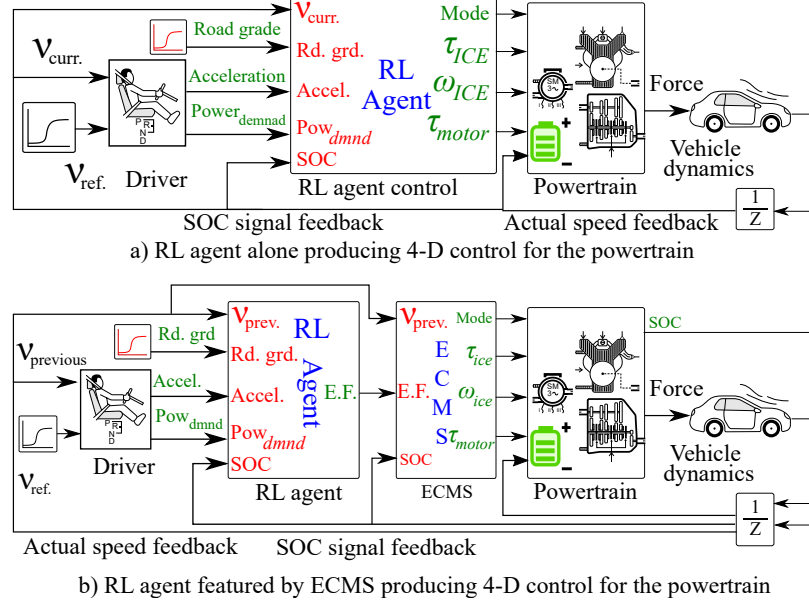


Figure 7.7: Conflation of RL and ECMS in online implementation of multi-mode EMS.

with a primary focus on fuel consumption minimization and conservation of charge-sustainability. ECMS strives to achieve these two objectives by activating different modes throughout the given drive cycle. Its equivalent operating cost governs the activation of a particular mode (amongst the three feasible modes for this powertrain configuration) compared to the same of the other two modes at every time-step.

The following equation gives the equivalent cost structure used by ECMS, where  $\dot{m}_{fuel}$  is the mass flow rate of fuel in (gm/sec),  $LHV$  is the lower heating value of fuel in (kJ/gm)

$$((J)_{equi.})|_{U,mode} = \dot{m}_{fuel} \times (LHV)_{fuel} + \lambda \times Pow_{batt}. \quad (7.7.2)$$

,  $\lambda$  is the equivalence factor between stored electrical energy in HVB and stored chemical energy in the fuel, and  $U$  represents the set of control actions corresponding to the particular mode. The set of control actions are delineated in the previous

**Algorithm 3:** Multi-mode ECMS

---

**Input :**  $\tau_{rqst_{drvr}}(t)$  &  $\omega_{rqst_{drvr}}(t)$  &  $\lambda(t)$   
**Output :**  $U(t)$ , where  $U = \{mode_{optim.}, \Psi_{mode_{optim.}}\}$ , where  
 $\Psi_{HEV} = \{\tau_{ice_{cmd}}, \tau_{mot_{cmd}}\}^T$ ,  
 $\Psi_{EV0} = \{\}, \Psi_{EV1} = \{\tau_{mot_{cmd}}\}$   
**Initialization :**  $\tau_{out} = \tau_{rqst_{drvr}}, \omega_{out} = \omega_{rqst_{drvr}}$   
**for**  $mode = i, i \in \{HEV, EV0, EV1\}$  **do**  
    sweep through all feasible combinations of  $\Psi_i$   
    **for each feasible combinations of  $\Psi_i$  do**  
        **compute:**  $\omega_{ice}, \tau_{ice}, \omega_{gen}, \tau_{gen}, \omega_{mot}$ ,  
        and  $\tau_{mot}$  corresponding to  $\Psi_i$   
        **compute:**  $\dot{m}_{fuel}$  (if  $i == HEV$ ),  $Pow_{batt}$ .  
         $J_{equi.i} = \dot{m}_{fuel} \times (LHV)_{fuel} + \lambda \times Pow_{batt}$ .  
     $\Psi_{i,optim.} = \arg \min_{\Psi_i} \|(J_{equi.i})_i\|$   
 $mode_{optim.} = \min(J_{equi.HEV}, J_{equi.EV0}, J_{equi.EV1})$   
 $U(t) = \{mode_{optim.}, \Psi_{mode_{optim.}}\}$

---

sub-section.

It has been reiterated on copious occasions in the literature that EF plays a crucial role in conserving charge-sustainability and optimality of the yielded control from ECMS. The value of EF ( $\lambda$ ) strictly governs the mode selection at any given time-step. EF's constant optimal value for the entire drive cycle can be obtained through the trial-and-error method or iteratively, given that the drive cycle is known in advance. However, prior knowledge of future driving conditions is not accurately available in most real-world driving scenarios. Hence, either rule-based control or PID controllers are employed to meet at least charge-sustenance criteria. However, the proportional, integral, derivative, and any other gains of the PID controller need manual tuning for different drive cycles. In this article, the authors propose an A3C agent for regulating the equivalence factor in real-world driving conditions with a cardinal objective of yielding near-optimal control in an unprecedented drive cycle, given that the agent is

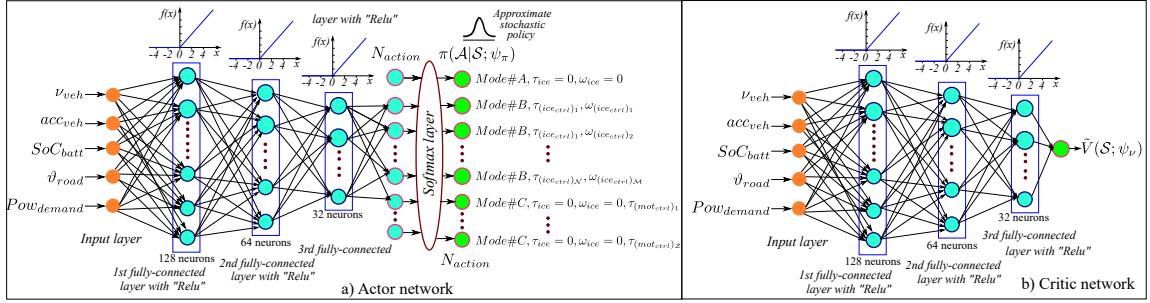


Figure 7.8: Deep neural network architecture for actor and critic network.

well acquainted with similar driving situations but not with the same drive cycle. The A3C agent, in the conflated version of the ECMS controller, decides the value of EF at each time-step. The equivalence factor emitted from the A3C agent is fed to the ECMS controller as an input, as shown in Fig.7.7.b. The EF action-space is comprised of 181 control actions and is quantized into  $EF \in \{1.3 : 0.1 : 2.3, \dots, 2.31 : 0.01 : 4\}$

### Articulation of DNN-based A3C Agent in Simulink®

The architectures of both actor and critic networks are depicted in Fig.7.8. Deep network designer tool from Mathworks® is employed to design both the networks. As depicted in Fig.7.8.a, the actor is articulated with an input layer (5 input ports), first hidden layer with “Relu” activation and 128 neurons, second hidden layer with “Relu” activation and 64 neurons, second hidden layer with “Relu” activation and 32 neurons, and finally, an output layer with *softmax* layer and  $N_{action}$  outputs.  $N_{action}$  is the number of discrete actions. As depicted in Fig.7.8.b, the critic network has the same architecture as actor network except the output layer. The critic network has only one output, i.e., the approximated state-value function corresponding to a certain state.

In the current endeavor, the number of hidden layers and neurons in the hidden

layers are selected through a trial-and-error method. Nevertheless, there are a few rule-of-thumb for choosing the number of neurons in a hidden layer, as enumerated here:

- The number of neurons in a hidden layer should fall between the size of output and input layer. For example, if the sizes of the output layer and input layer are one hundred and ten, respectively, the number of neurons in the hidden layer can be in between ten and one hundred.
- $N_{neuron} = \frac{2}{3} * N_{input} + N_{output}$ , where  $N_{neuron}$ ,  $N_{input}$ , and  $N_{output}$  are the number of neurons in a hidden layer, size of the input layer, and size of output layer, respectively.
- $N_{neuron} = \sqrt{N_{input} \times N_{output}}$

For a decent complex problem having a large feature’s pool can be started with two to three hidden layers. Learn rate ( $\alpha$ ) refers to the rate of change of weights of the network. Discount factor ( $\gamma$ ) regulates the contribution of future rewards in the computation of return from the current step. Number of step to look-ahead refers to “ $N$ ” as shown in Fig.7.3.

Convergence time is inversely proportional to the learning rate ( $\alpha$ ). Using a minimal value ( $\alpha \leq 10^{-5}$ ) of  $\alpha$  will procrastinate the convergence of the A3C agent. However, a reasonably small value, i.e., ( $\approx 0.001$ ), is a wise choice for  $\alpha$  since it will not let the agent converge to a sub-optimal solution. A higher  $\alpha$  value will destabilize the training process and eventually diverge the agent from the optimal solution.

In [178], the authors introduced the concept of “*entropy maximization*” to discourage the agent from selecting only one particular control action among a finite set of actions

all of which give similar reward. According to [178] and [173], *entropy function* ( $H$ ) of policy is expressed in the following:

$$H = -\mathcal{E} \sum_{k=1}^M \pi_k(S_t; \psi_\pi) \ln \pi_k(S_t; \psi_\pi) \quad (7.7.3)$$

, where  $\mathcal{E}$  is entropy loss-weight,  $\pi_k(S_t; \psi_\pi)$  is the probability of choosing an action  $A_k$  in state  $S_t$ , and  $M$  is the number of maximum possible actions.

It is likely to happen during the update of both actor and critic weights that their gradient can increase exponentially in magnitude. This phenomenon is widely known as the “*gradient explosion*”. *Gradient explosion* can engender instability in both networks’ training, and it can diverge soon after the commencement of training. Gradient clipping is a well-known wield to circumvent such a problem. If the gradient at any time-step exceeds the designer’s threshold value, a designated gradient clipping method will be enabled to restrict the gradient value. Entropy loss weight ( $\mathcal{E}$ ) refers to the co-efficient regulating the strength of entropy regularization. “L2-Regularization” is employed in both the networks to prevent them from overfitting. *Gradient threshold* refers to a precautionary threshold limit for gradient update. *Gradient threshold method* refers to a method used to clip gradient values that exceed the *gradient threshold*. A few of the available methods are “*L2norm*”, “*Global-L2norm*”, “*Absolute value*”. *Optimizer* refers to the optimization algorithm used to update weights of the networks iteratively based on training data. A few of the available optimization algorithms are *stochastic gradient descent (SGD)*, *SGD with momentum (SGDM)*, *RMSProp*, *Adam*. Tab.7.4 furnishes the list of hyper-parameters used for A3C agent in this chapter. In the current research, the hyperparameters are tuned manually, which is time consuming and sometime can be exasperating. Self-tuning

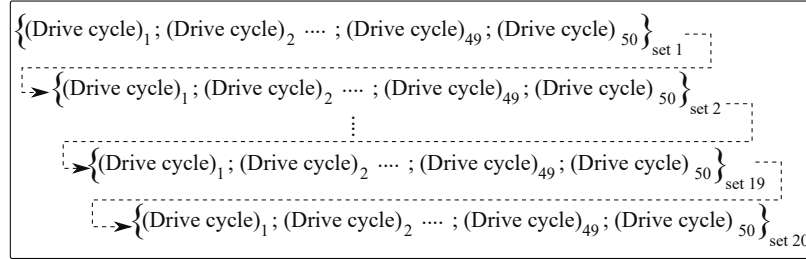


Figure 7.9: Sequence of drive cycles for training schedule.

DRL such as self-tuning actor-critic (STAC) employs a meta-gradient approach to tune the hyperparameters online with 30% additional computational burden and find an optimal trade-off among the performance metrics of the actor-critic [179].

## 7.8 Results and Discussion

The first and foremost required thing is data of real-world drive cycles to train the A3C agent for real-world driving scenarios and test its performance in similar situations. A handful of drive cycles have been acquired with an in-house data acquisition system at McMaster Automotive Resource Center (MARC). However, an A3C agent must be facilitated with many more drive cycles than just a handful of drive cycles. A Markov chain model (MCM) can be employed to generate a copious number of random drive cycles with the acquired driving data characteristics. Fig.7.12 shows a few of the acquired drive cycles and a few of the randomly generated drive cycle (RGDC)s. The MCM-based random drive cycle generator tool features keeping the cycles' length the same for every generated one.

### 7.8.1 Random Drive-cycle Generation with Markov Chain Model

A Markov chain model (MCM) of the power demand has been successfully developed to generate numerous random drive cycles. Data of 58 previously collected drive cycles have been used to articulate the MCM. Random drive cycles can be generated using the MCM, but the drive cycles will carry the same transition probabilities of power demand extracted from the collected drive cycles' ensemble. The utilization of random drive cycle generation with MCM has been proposed in section 6.3.3. The MCM models the stochastic behavior of specific parameters' variation in any time-series distribution. In this research study, power demand is chosen to develop the MCM using the collected drive cycle data. The MCM of the vehicle's power demand represents the trend of driving behavior throughout the ensemble in the form of a probability matrix. Every newly generated drive cycle will be different, but their stochastic driving behavior will be the same because they are generated from the same MCM. The MCM generates power demand of next time-step via forecasting with the aid of the following transition probability equation:

$$Pr(Pow_{dem_{k+1}} = Pow_{dem}^j | Pow_{dem_k} = Pow_{dem}^i, v_k = v^l) = Pr_{il,j} \quad (7.8.1)$$

The above-mentioned probability represents the probability of transitioning to  $Pow_{dem}^j$  in the next time-step if the power demand and speed of vehicle are  $Pow_{dem}^i$  and  $v^l$  at the current time-step.

The probabilities are stored in a matrix referred to as the transition probability matrix (TPM). First of all, the ranges of power demand values and velocity values across



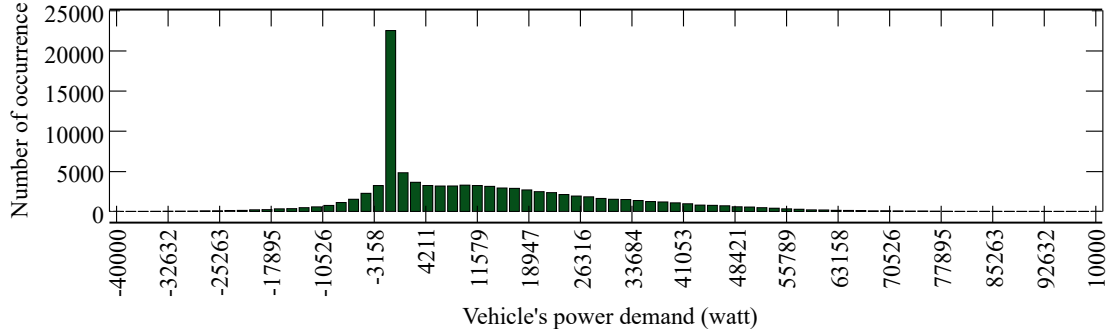


Figure 7.10: Range of vehicle's power demand over 58 real-world drive cycles.

the ensemble of drive cycles have to be discretized into a pre-defined resolution like  $\{Pow_{dem}^1, Pow_{dem}^2, \dots, Pow_{dem}^{N_p}\}$  and  $\{vel^1, vel^2, \dots, vel^{N_m}\}$ .

Such a discretization have been done non-linearly based on observation of the power demand distribution throughout all collected data point. Since the collected data does not contain the vehicle's power demand, it has been calculated with the following equation and vehicle specification.

$$\begin{aligned}
 Power_{demand} &= \tau_{transmission\_out} \times \omega_{transmission\_out} \\
 &= \tau_{final\_drive\_in} \times \omega_{final\_drive\_in} \\
 &= (Final\_drive\_ratio) \times \tau_{wheel} \times \frac{\omega_{wheel}}{(Final\_drive\_ratio)}
 \end{aligned} \tag{7.8.2}$$

, where  $\tau_{wheel} = (\text{torque required to counter inertial load} + \text{torque required to counter rolling resistance from road} + \text{torque required to counter aerodynamic force} + \text{torque required to counter grade load from road slope})$ . Although grade force and an aerodynamic force acts on the vehicle body, these two forces can be converted into equivalent wheel torque with a one-dimensional (1-D) approximation of the whole vehicle body.

Using Eq.7.8.2, vehicle's power demand has been calculated for 58 real-world drive

cycles and shown in Fig.7.10. The power demand range and speed range are discretized into 106 values and 51 values respectively. State of the MCM is comprised of all combinations of power demand values and speed values. Hence, total feasible number of state is 5406. Now, it is to be checked that how many unique states are sampled throughout the collected drive cycles. Using Nearest-neighbor quantization the sequence of collected data ( $Pow_{dem}, vel$ ) is mapped into a sequence of discretized states ( $Pow_{dem}^i, vel^j$ ). 3915 numbers of unique states have been found and hence, a *TPM* matrix is initialized with a size of  $3915 \times 3915$ .

Transition probability is estimated using maximum likelihood estimator which counts collected data as:

$$(TP)_{ij} = \frac{Total\_number\_transitions\_from\_state\_i\_to\_state\_j}{Total\_number\_transitions\_from\_state\_i\_to\_all\_states} \quad (7.8.3)$$

The Markov chain model (MCM) can be generated for the above mentioned TPM using "*dtmc*" function in MATLAB® with just a simple command " $TPM = dtmc(TPM)$ ".

The whole process of generating new drive cycles of any specific duration is depicted in Fig. 7.11. Both speed profile and power demand profile can be generated from the MCM. While generating drive cycles with "*simulate*" function in MATLAB®, the initial state can be specified by the user. This feature is very beneficial from the perspective of onboard RL implementation for EMS as described in 6.3.3.

Initially, only fifty RGDCs are utilized for training the A3C agent, and another RGDC is used to testify the performance of the trained A3C agent as an EMS controller for the electrified powertrain.

The set of fifty drive cycles is repeated twenty times during the training phase so that the agent gets trained for 1000 episodes. For keeping the training phase simple,

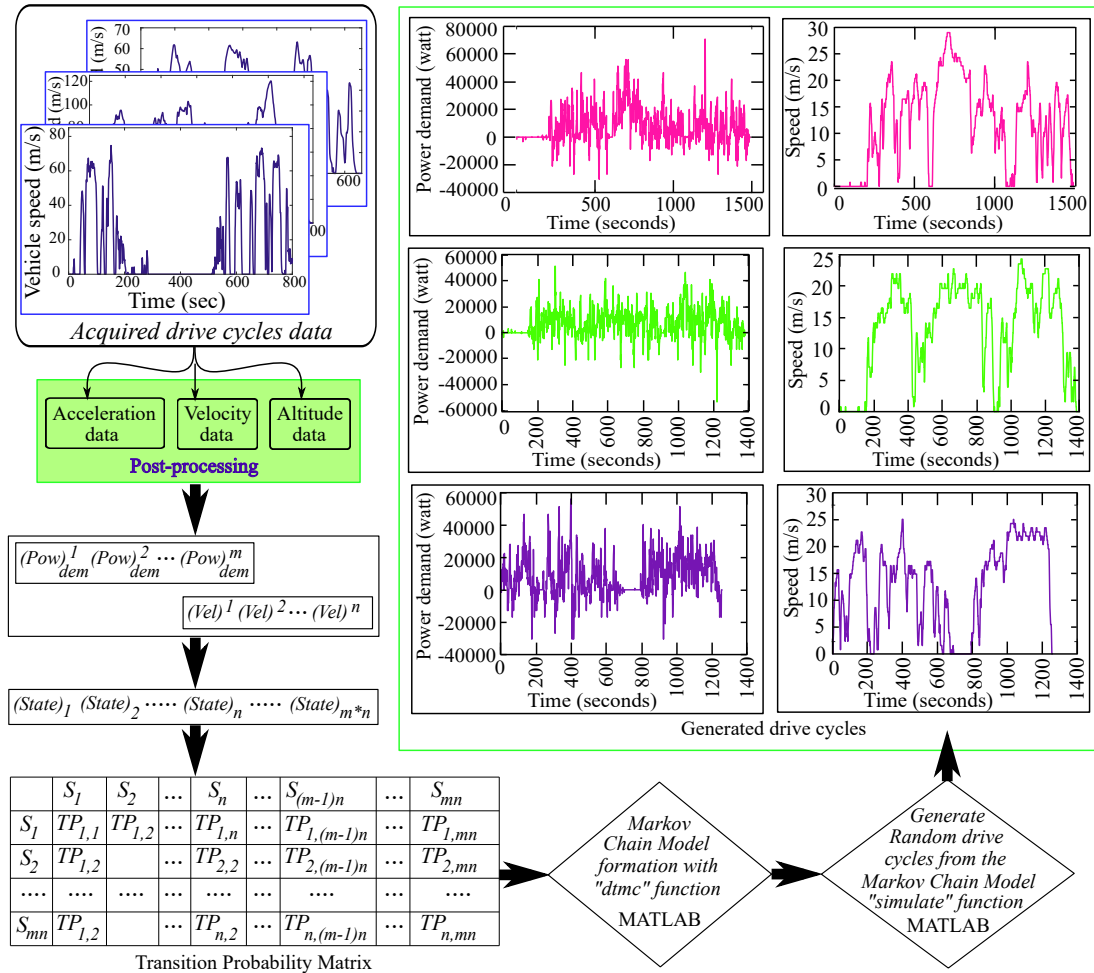


Figure 7.11: Random drive cycle generation with Markov chain model.

the order of RGDCs is kept unaltered in every set as depicted in Fig.7.9. As shown in Fig.7.13, the three cardinal components of training history are episodic return, average return, and episodic  $Q_0$  value. Episodic return is the cumulative cost in each episode. The "episodic returns" are averaged out after every ten episodes, and the average return value represents that. The  $Q_0$  value corresponding to each episode refers to the agent's estimation of optimal cumulative return at the very first time-step of each episode. The agent's performance is testified after 1000 training episodes

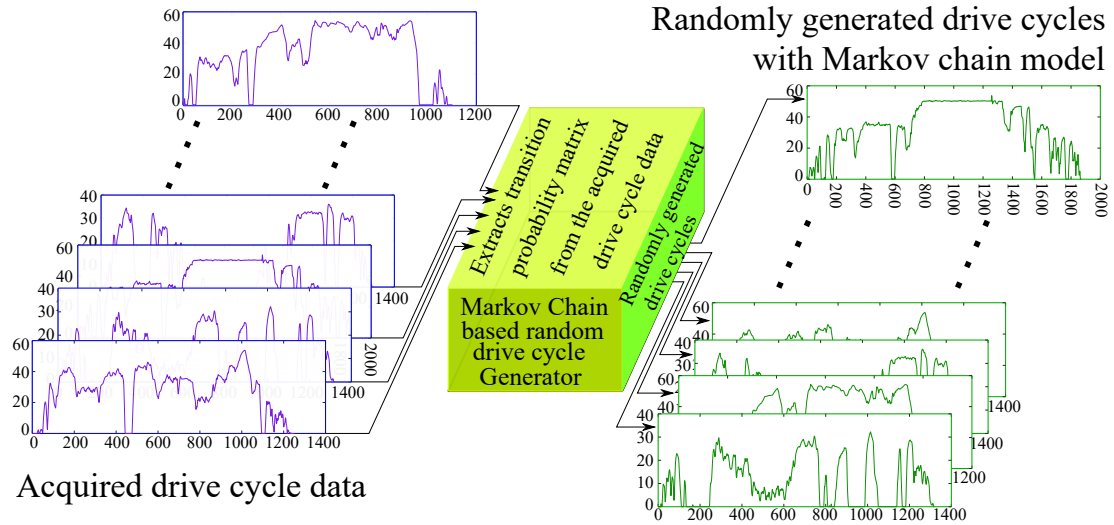


Figure 7.12: Acquired and randomly generated drive cycles used for training.

by employing it as an EMS controller for a drive cycle randomly generated with the MCM. The rudimentary expectation from the trained agent is its capability of satisfying the charge sustenance. If the trained A3C agent can suffice the charge sustenance at the end of the randomly generated drive cycle, it would indicate that the agent will most probably engender charge sustainability for any RGDC generated by the MCM. Fig.7.14.a corroborates the A3C agent’s charge sustaining competency, which starts with 40% battery SOC and finishes the drive cycle with 40.5% battery SOC value. Fig.7.14.b presents the time-series profiles of powertrain mode and fuel consumption throughout the drive cycle. Although the powertrain can exhibit three modes, A3C agent-based EMS has never activated the two-motor all-electric mode throughout the RGDC.

Fig.7.15 depicts the time-series of ICE speed and generator speed with an enlarged view at the top right corner. A medium-fidelity ICE model with cranking and engine-braking dynamics is used instead of a steady-state ICE model in the FSM-based

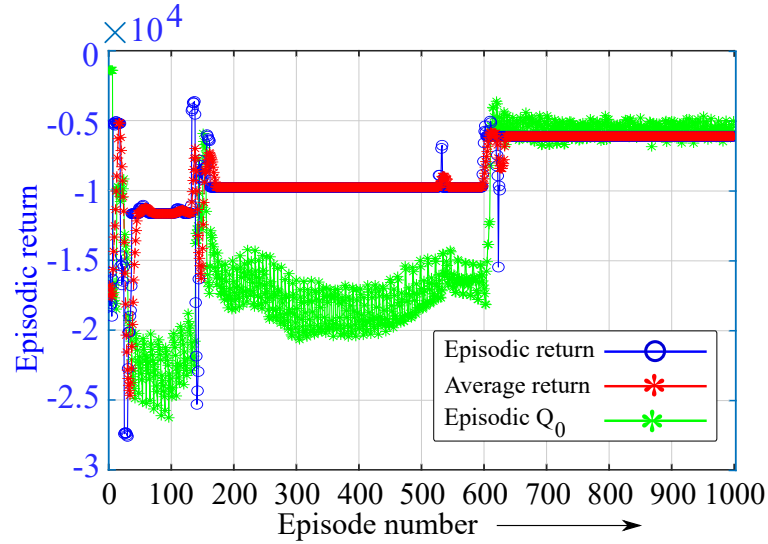


Figure 7.13: Training history of A3C agent-based EMS controller through 1000 episodes.

powertrain model. The stark difference between requested ICE-speed and actual ICE-speed, as shown at the left bottom corner of Fig.7.15, reveals the fact that a medium-fidelity ICE cannot achieve the requested speed as instantaneously as it receives the command from the EMS controller due to inertia and other physical lags. Fig.7.16 presents ICE torque and generator torque time-series, including the associated torque dynamics during ICE-cranking and ICE-braking. The generator exerts positive torque to crank the ICE until ICE starts firing and producing useful torque. The drive cycle's overall fuel consumption with the trained A3C agent as the EMS controller is 959 grams. It is time to compare the above analyzed A3C agent-based EMS controller's proficiency with a benchmark real-time EMS controller. The authors in this article have chosen optimal ECMS as the benchmark controller. The optimal equivalence factor of ECMS for the same RGDC used to testify the A3C agent is obtained through the trial-and-error method.

Fig.7.17, Fig.7.18, and Fig.7.19 present the comparison study between performance

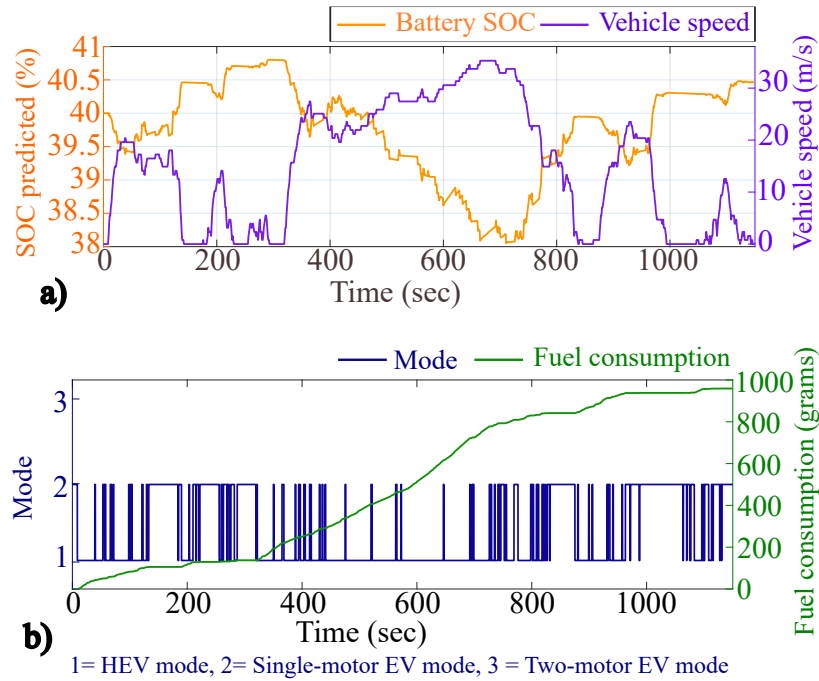


Figure 7.14: a) Randomly generated drive cycle, which is used to testify the A3C agent, plotted with right axis. Battery SOC variation throughout the drive cycle is plotted with left axis. b) Time-series of powertrain mode and fuel consumption are plotted with left axis and right axis respectively.

metrics of ECMS and A3C agent. While charge sustenance criteria and fuel consumption are chosen as the cardinal metrics for conducting the comparison study, mode-shift frequency is an indirect measurement of drivability. Fig.7.17 elucidates that the A3C agent yields a lesser frequency of mode-shifts compared to ECMS. One of the heavily crowded regions of the mode profile is enlarged for ECMS and A3C agents to compare their drivability performance. The stark difference in the enlarged section of ECMS and A3C agent corroborates that the A3C agent is better in drivability than ECMS. However, there is still scope for improvement in drivability, and it can be implemented by adding a penalty for too frequent mode-shifts in the *immediate consequence function*.

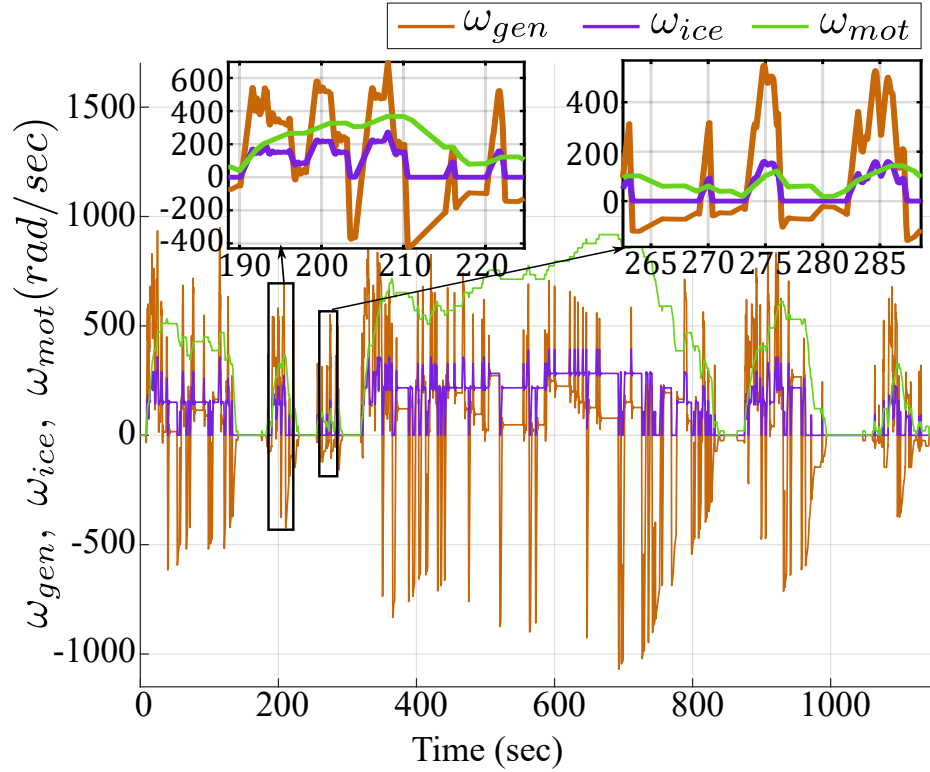


Figure 7.15: Time-series of  $\omega_{ice}$ ,  $\omega_{gen}$ , and  $\omega_{mot}$  throughout the real-world drive cycle.

Both optimal ECMS and A3C agent can suffice the charge sustenance criteria. However, the optimal ECMS knew the drive cycle information in advance, whereas the A3C agent did not know the drive cycle in advance. Leveraging on the apriori knowledge of drive cycle, the optimal ECMS can achieve perfect charge-sustainability, but the equivalence factor is modified so that the end SOC value reaches as close as that obtained by the A3C agent. The ECMS is designed in this way only to present a fair comparison of fuel consumption value between the two EMS controllers, as shown in Fig.7.19. Both the controllers approximately finish the drive cycle at 40.5% of battery SOC.

The study on the effectiveness of the A3C agent over optimal ECMS as an EMS

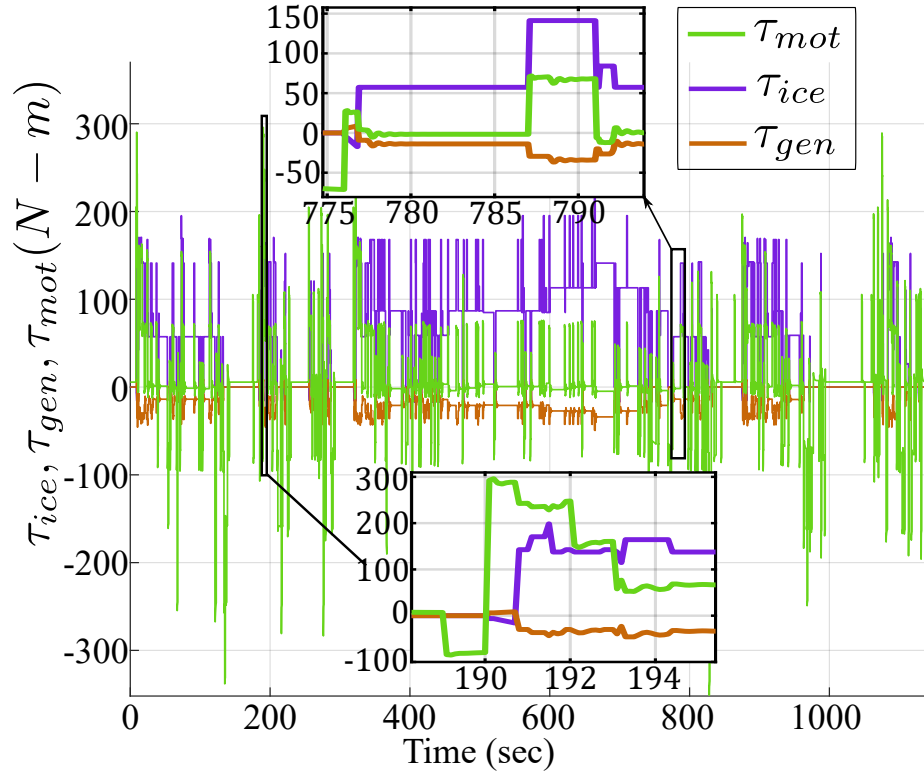


Figure 7.16: Time-series of  $\tau_{ice}$ ,  $\tau_{gen}$ , and  $\tau_{mot}$  throughout the real-world drive cycle.

controller for ePT culminates in comparing overall fuel consumption between two controllers for a given random drive cycle. As furnished in Fig.7.19, the comparison corroborates the superiority of the A3C agent over optimal ECMS in fuel economy. While optimal ECMS concedes 978 grams of fuel, A3C agent concedes 959 grams of fuel, which is 1.9% lesser consumption than optimal ECMS.

No matter how well the A3C agent performs for a real-world drive cycle, its prowess is not recognized until proven for a few of the federal drive cycles. Hence, the A3C agent, trained with 1000 real-world drive cycles, is tested on two federal drive cycles, i.e., UDDS and WLTC, to validate its lucrative features.

Left axis of Fig.7.20 and Fig.7.21 depict the mode profile and prime-movers' torque, respectively controlled by A3C agent. Fig.7.22 and left axis of Fig.7.23 depict the



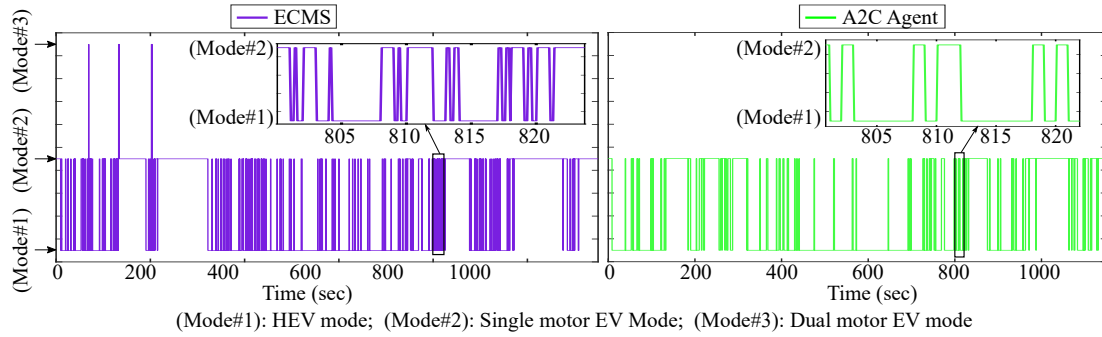


Figure 7.17: Comparison of powertrain mode profiles between ECMS and A3C agent when applied on real-world drive cycle.

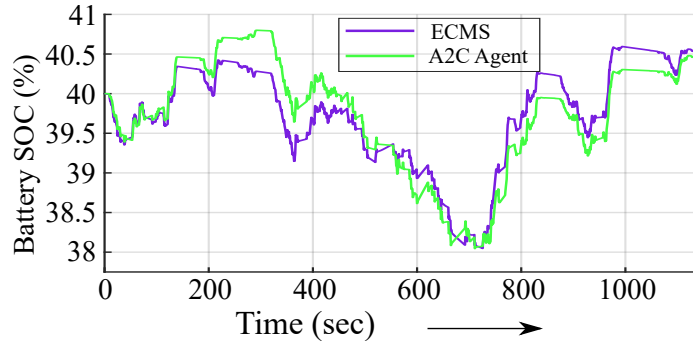


Figure 7.18: Comparison of battery SOC profiles between ECMS and A3C agent when applied on real-world drive cycle.

prime-movers' angular speed profile and variation of battery SOC, respectively as a result of the A3C-based EMS control. Right axis of Fig.7.20 shows the time-series of fuel consumption across the UDDS. The A3C agent has reaffirmed its capability of finishing the UDDS well-within the permissible limit of SOC deviation, as shown in the Fig.7.23. The SOC-deviation is just 0.4% which is very close to absolute charge-sustainability.

As mentioned before, the online framework comprised of training and testing of the

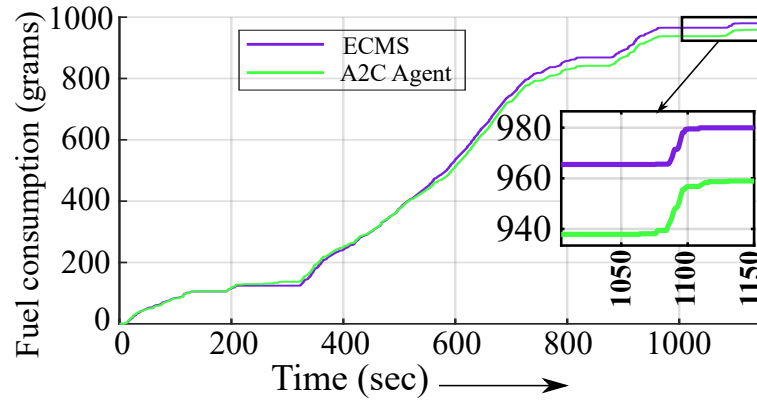


Figure 7.19: Comparison of fuel consumption between ECMS and A3C agent when applied on real-world drive cycle.

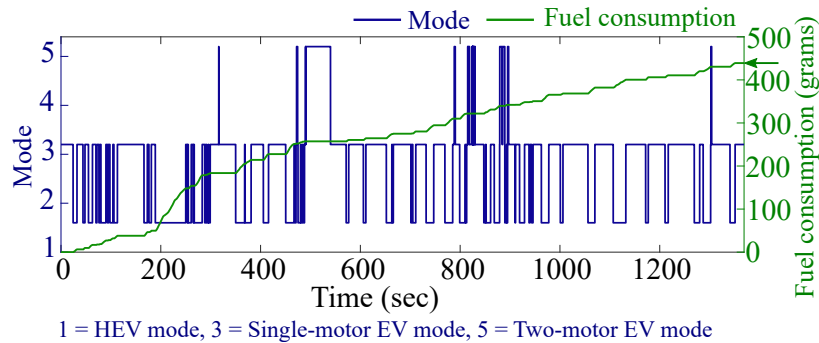


Figure 7.20: Time-series of powertrain mode obtained for UDDS with A3C agent as an EMS control strategy.

A3C agent mimics real-time interaction between the A3C agent and an ePT in unfamiliar real-world driving scenarios. Suppose, an HEV is equipped with a data acquisition system, the A3C-based EMS mentioned above, and the Markov chain-based random drive cycle generator. The current driving data is acquired periodically and fifty random drive cycles can be generated through the Markov chain model based on the recently acquired driving data. Now, the A3C agent can be trained with these fifty RGDCs. The training of the A3C agent continues in the background, and there is no compulsion that training has to complete within a single discrete time-step.

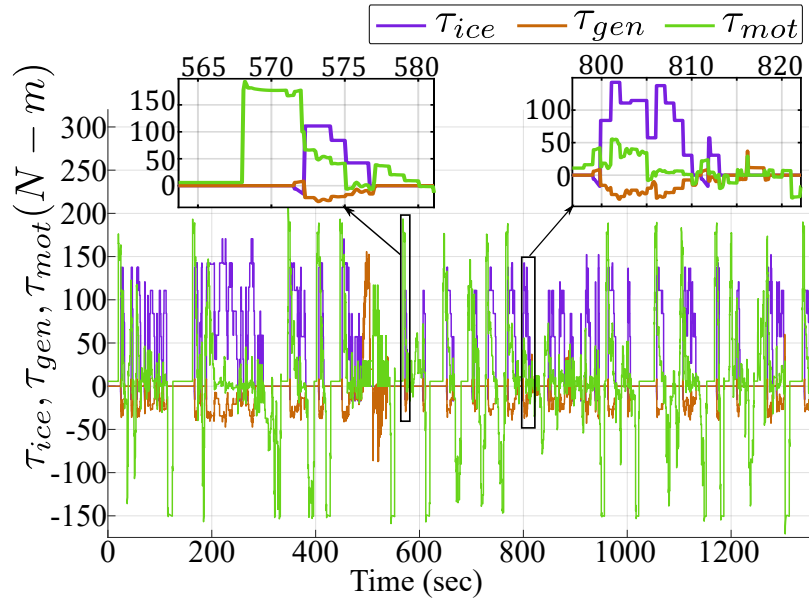


Figure 7.21: Time-series of  $\tau_{ice}$ ,  $\tau_{gen}$ , and  $\tau_{mot}$  obtained for UDDS with A3C agent as an EMS control strategy.

However, a brief sensitivity analysis, as depicted in Fig.7.24 and Fig.7.25, is required to inspect the effect of training iterations on the performance of the A3C agent.

In this article, ICE-activation's total number throughout the drive cycle is chosen as the drivability metric. The radar plots for three different testing drive cycles in Fig.7.25 depicts the different level of dexterity of the A3C agent after different training level. Conspicuously, the A3C agent has significantly alleviated the drivability score at every trained level.

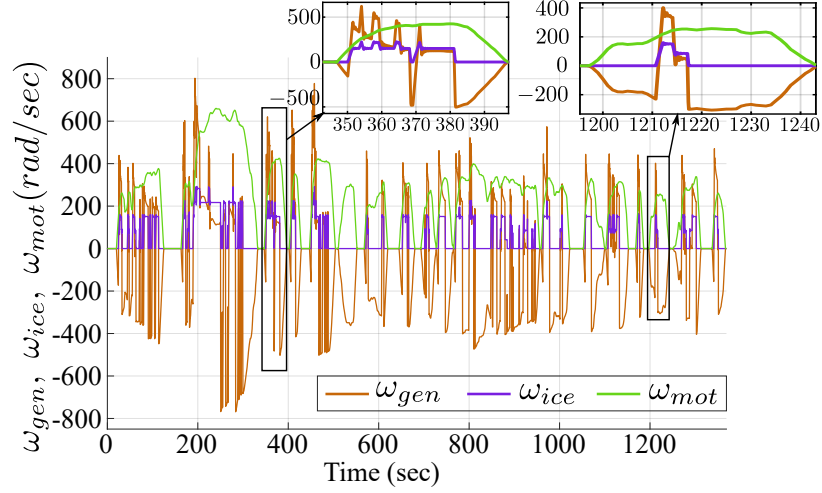


Figure 7.22: Time-series of  $\omega_{ice}$ ,  $\omega_{gen}$ , and  $\omega_{mot}$  obtained for UDDS with A3C agent as an EMS control strategy.

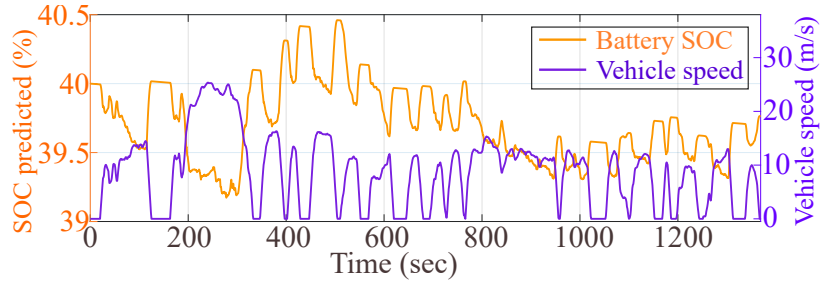


Figure 7.23: Time-series of battery SOC variation throughout UDDS with A3C agent as an EMS control strategy.

Table 7.1: Upper and lower bounds of state variables

State variable	Unit	Upper bound	Lower bound
Vehicle's speed	m/s	50	0
Vehicle's acceleration	m/s <sup>2</sup>	+10	-10
SOC of HVB	[%]	100	0
Driver's torque demand	kWatt	150	-100
Road grade	[%]	30	-30

Table 7.2: Feasible combinations Control variables of different modes

Feasibility	Control Variables				Fulfilling criteria	Generator torque
	Mode ctrl.	$\omega_{ice}$ ctrl.	$\tau_{ice}$ ctrl.	$\tau_{mot}$ ctrl.		
Feasible	EV#1	==0	==0	$\neq 0$	$\tau_{mot} = \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Infeasible	EV#1	$\neq 0$	==0	$\neq 0$	$\tau_{mot} = \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Infeasible	EV#1	==0	$\neq 0$	$\neq 0$	$\tau_{mot} = \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Infeasible	EV#1	==0	==0	$\neq 0$	$\tau_{mot} \neq \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Feasible	HEV	$\neq 0$	$\neq 0$	$\neq 0$	$\tau_{mot} = \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Feasible	HEV	$\neq 0$	==0	$\neq 0$	$\tau_{mot} = \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Infeasible	HEV	==0	$\neq 0$	$\neq 0$	$\tau_{mot} = \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Infeasible	HEV	$\neq 0$	$\neq 0$	$\neq 0$	$\tau_{mot} \neq \frac{1}{\beta_2} \tau_{out} - \frac{\beta_1}{\beta_2 * (1 + \beta_1)} \tau_{ice}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Feasible or infeasible	HEV	$\neq 0$	$\neq 0$	$\neq 0$	Feasible if $\tau_{ice} = \frac{(1 + \beta_1)}{\beta_1} \tau_{out}$	$\tau_{gen} = -\frac{1}{1 + \beta_1} \tau_{ice}$
Infeasible	EV#2	$\neq 0$	==0	$\neq 0$	$\tau_{gen} = \left( \frac{\beta_2 * \tau_{mot} - \tau_{out}}{\beta_1} \right)$	
Infeasible	EV#2	==0	$\neq 0$	$\neq 0$	$\tau_{gen} = \left( \frac{\beta_2 * \tau_{mot} - \tau_{out}}{\beta_1} \right)$	
Feasible	EV#2	==0	==0	$\neq 0$	$\tau_{gen} = \left( \frac{\beta_2 * \tau_{mot} - \tau_{out}}{\beta_1} \right)$	

Table 7.3: Samples of feasible and infeasible combinations of 4-D action variables

Given condition	4-D control: discrete action-space for the A3C agent				
$\tau_{out}$ (N-m)	Mode	$\tau_{ice}$ (N-m)	$\omega_{ice}$ (rad/s)	$\tau_{mot.}$ (N-m)	Feasi./ Infeasi.
150	EV#1	0	0	60	Feasible
	EV#1	125	0	60	Infeasible
	EV#1	0	250	60	Infeasible
	EV#1	70	225	0	Infeasible
150	HEV	120	150	24.4	Feasible
	HEV	120	150	20	Infeasible
	HEV	190	160	40	Infeasible
	HEV	100	130	30.336	Feasible
	HEV	130	225	21.44	Feasible
150	EV#2	0	0	50	Feasible
	EV#2	0	0	60	Feasible
	EV#2	0	0	30	Feasible
	EV#2	60	120	any value	Infeasible

Table 7.4: Hyper-parameters for implementing asynchronous actor-critic agent in Simulink®

Name of Hyper-parameters	Range of value	Used value:
Learn rate( $\alpha$ )	$1 \geq \alpha \geq 10^{-5}$	0.001
Discount factor( $\gamma$ )	$1 \geq \gamma \geq 0$	0.995
Num. of step to look-ahead	$T_{max} \geq N \geq 1$	$T_{max}$
Entropy Loss Weight( $\mathcal{E}$ )	$1 \geq \mathcal{E} \geq 0$	$\mathcal{E} = 0.2$
Gradient threshold ( $\mathcal{G}$ )	$\infty \geq \mathcal{G} \geq 0$	3.5
Gradient threshold method	"L2norm"	
Optimizer	"Adam"	

Table 7.5: Definition of Transition Probability Matrix

—	$Pow_1, vel_1$	$Pow_1, vel_2$	...	$Pow_1, vel_n$	...	$Pow_m, vel_1$	...	$Pow_m, vel_n$
—	$S_1$	$S_2$	...	$S_n$	...	$S_{(m-1)*n}$	...	$S_{m*n}$
$S_1$	$TP_{1,1}$	$TP_{1,2}$	...	$TP_{1,n}$	...	$TP_{1,(m-1)n}$	...	$TP_{1,m*n}$
$S_2$	$TP_{2,1}$	$TP_{2,2}$	...	$TP_{2,n}$	...	$TP_{2,(m-1)n}$	...	$TP_{2,m*n}$
...	...	...	...	...	...	...	...	...
$S_n$	$TP_{n,1}$	$TP_{n,2}$	...	$TP_{n,n}$	...	$TP_{n,(m-1)n}$	...	$TP_{n,m*n}$
...	...	...	...	...	...	...	...	...
$S_{m*n}$	$TP_{m*n,1}$	$TP_{m*n,2}$	...	$TP_{m*n,n}$	...	$TP_{m*n,(m-1)n}$	...	$TP_{m*n,m*n}$

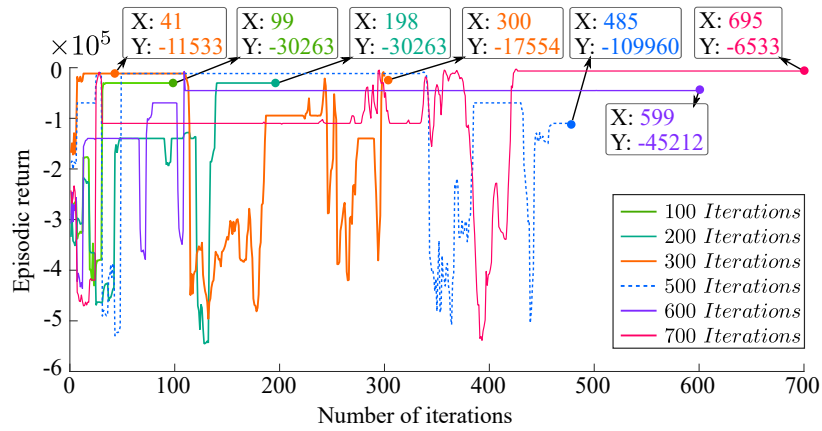


Figure 7.24: Effect of training iteration numbers on the agent’s convergence to near-global optimality.

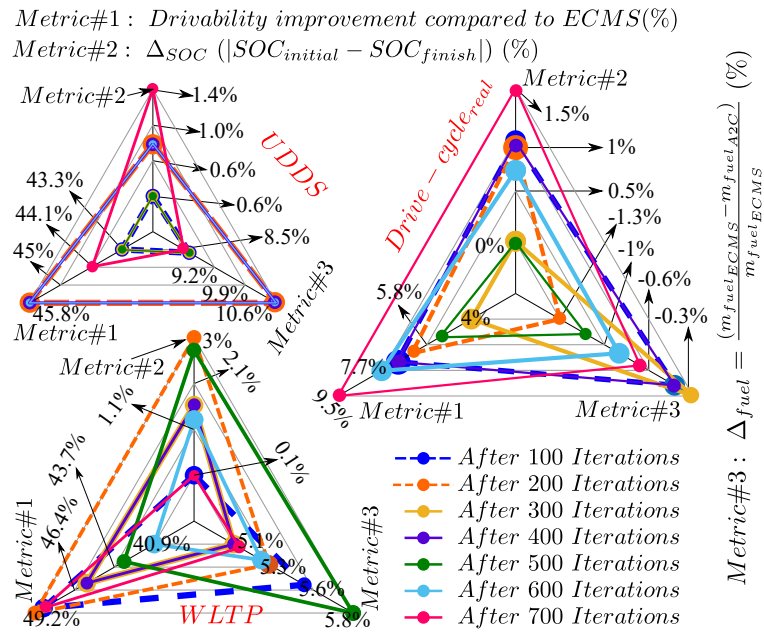


Figure 7.25: Effect of training iteration numbers on the agent’s performance (drivability, charge-sustenance, and fuel-consumption) compared to optimal ECMS.



Table 7.6: Effect of training iteration numbers on the agent’s performance (drivability, charge-sustenance, and fuel-consumption) compared to optimal ECMS

Drive cycle	UDDS			Real-world			WLTP		
$Iter^n$	$SOC_{RL_{end}}$	$\Delta m_{fuel}$	$\Delta activa_{ice}^n$	$SOC_{RL_{end}}$	$\Delta m_{fuel}$	$\Delta activa_{ice}^n$	$SOC_{RL_{end}}$	$\Delta m_{fuel}$	$\Delta activa_{ice}^n$
	$\Delta SOC_{end}$	$m_{fuel_{RL}}$	$activa_{RL}^n$	$\Delta SOC_{end}$	$m_{fuel_{RL}}$	$activa_{RL}^n$	$\Delta SOC_{end}$	$m_{fuel_{RL}}$	$activa_{RL}^n$
	RL agent	$m_{fuel_{ecms}}$	$activa_{ecms}^n$	RL agent	$m_{fuel_{ecms}}$	$activa_{ecms}^n$	RL agent	$m_{fuel_{ecms}}$	$activa_{ecms}^n$
100	39.765%	RL 439	RL 55	41%	RL 591.5	RL 145	40.145%	RL 1063	RL 55
	0.235%	ecms 480	ecms 97	1%	ecms 589	ecms 156	0.145%	ecms 1126	ecms 107
200	40.815%	RL 479.2	RL 58	40.94%	RL 588.9	RL 145	43%	RL 1168	RL 63
	0.815%	ecms 536	ecms 107	0.94%	ecms 581	ecms 155	3%	ecms 1233	ecms 124
300	39.765%	RL 439	RL 55	39.98%	RL 551.7	RL 144	41.6%	RL 1115.6	RL 62
	0.235%	ecms 480.5	ecms 97	0.02%	ecms 550.2	ecms 150	1.6%	ecms 1175	ecms 114
400	40.815%	RL 479.2	RL 58	41%	RL 591.5	RL 145	41.6%	RL 1115.6	RL 62
	0.815%	ecms 536	ecms 107	1%	ecms 589.1	ecms 156	1.6%	ecms 1175	ecms 114
500	39.765%	RL 439	RL 55	39.99%	RL 557.3	RL 142	37.27%	RL 965.1	RL 53
	0.235%	ecms 480.5	ecms 97	0.01%	ecms 551	ecms 150	2.73%	ecms 1025	ecms 93
600	40.815%	RL 479.159	RL 58	40.716%	RL 582.2	RL 144	38.7%	RL 1015	RL 55
	0.815%	ecms 536	ecms 97	0.716%	ecms 577.2	ecms 156	1.3%	ecms 1071	ecms 93
700	38.58%	RL 395.34	RL 52	41.52%	RL 609	RL 143	40.146%	RL 1063.8	RL 55
	1.42%	ecms 432	ecms 93	1.52%	ecms 604.8	ecms 158	0.146%	ecms 1121	ecms 107

## 7.9 summary

The original contributions of this article are reiterated here with brevity. A detailed review of the functional approximation based implementation of RL-agents is presented here. The review has not only just compared the existing functional approximation-based implementations, but their evolution, the intertwined aspects among themselves, and the essential stages of every implementation are also delineated with substantial importance. Asynchronous actor-critic is one of the most dexterous DNN-based RL agents among all the contemporary ones in the literature, and hence, it is selected as the real-time EMS controller for this study. One of this article's prime focuses is to emulate the real-world interaction between an RL agent-based EMS controller and an actual HEV. For such an emulation in a MIL platform, both the vehicle and the RL agent should be online implementable and compatible with each other. Hence, both vehicle and the agent models are developed in the Simulink® environment. The study presented here corroborates that the A3C agent can yield near-global optimal performance for unfamiliar drive cycles (standard federal or real-world drive cycles) in the future if it is trained with RGDCs generated based on current driving data.

Extra attention has been paid while modeling the powertrain dynamics to take the dynamics associated with mode-shifts, ICE cranking, and ICE braking into account. Not only the mathematical foundation of A3C agent but its implementation in Simulink® is also well illustrated. Such an illustration will insinuate the future researchers that a single RL agent's employment as a direct EMS controller can be a little convoluted, especially for multi-mode ePT where different modes have different control variables. Finally, a version of the A3C agent, conflated with conventional

ECMS, is articulated for the multi-mode ePT. The effectiveness of such a conflated version of the A3C agent is proved to be performing slightly better than optimal ECMS for a drive cycle that is randomly generated with MCM. In a nutshell, this article facilitates a comprehensive picture of DNN-based implementation of A3C agent, which has the potential of being a real-time and near-global optimal controller for real-world driving situations.

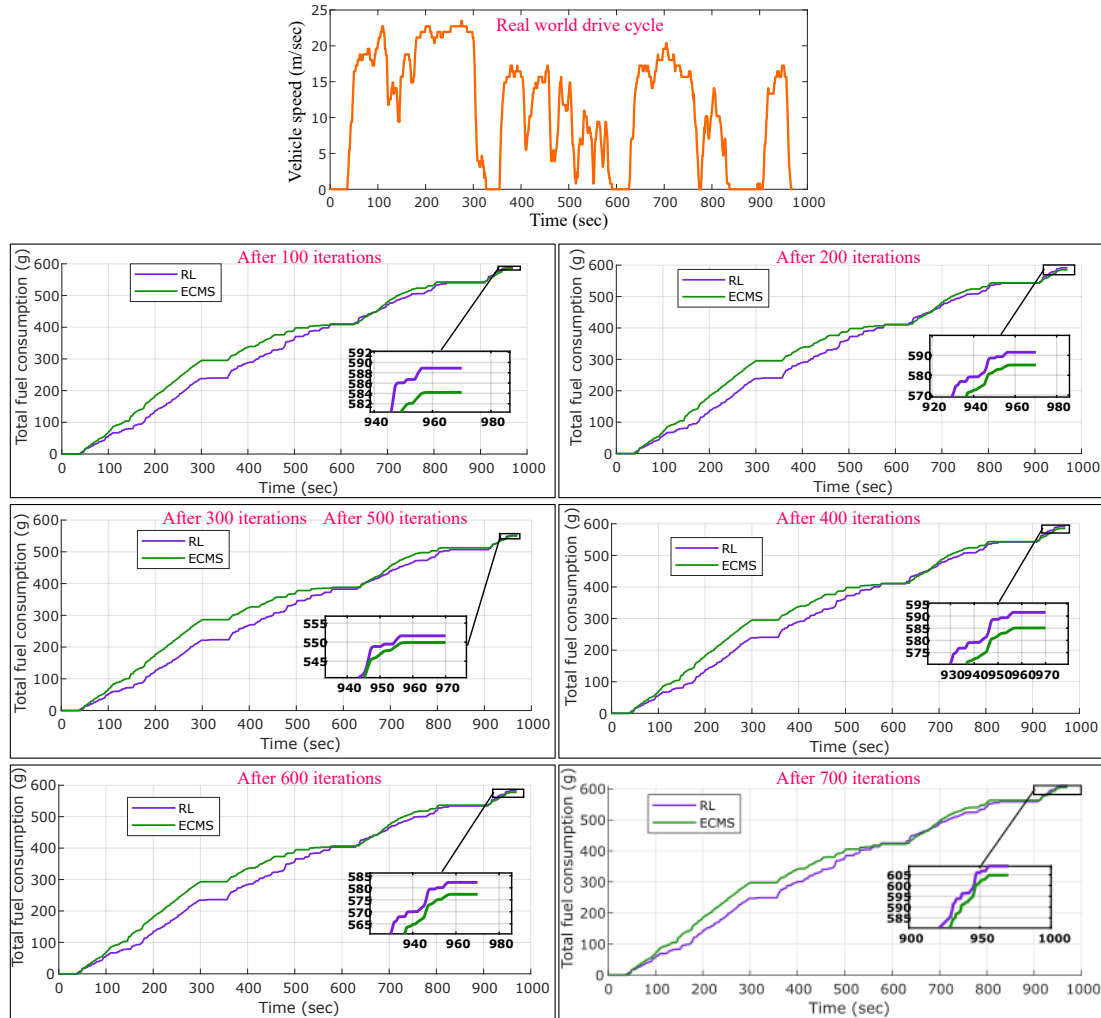


Figure 7.26: Effect of training iteration numbers on the agent’s fuel consumption performance compared to optimal ECMS for real-world drive cycle.

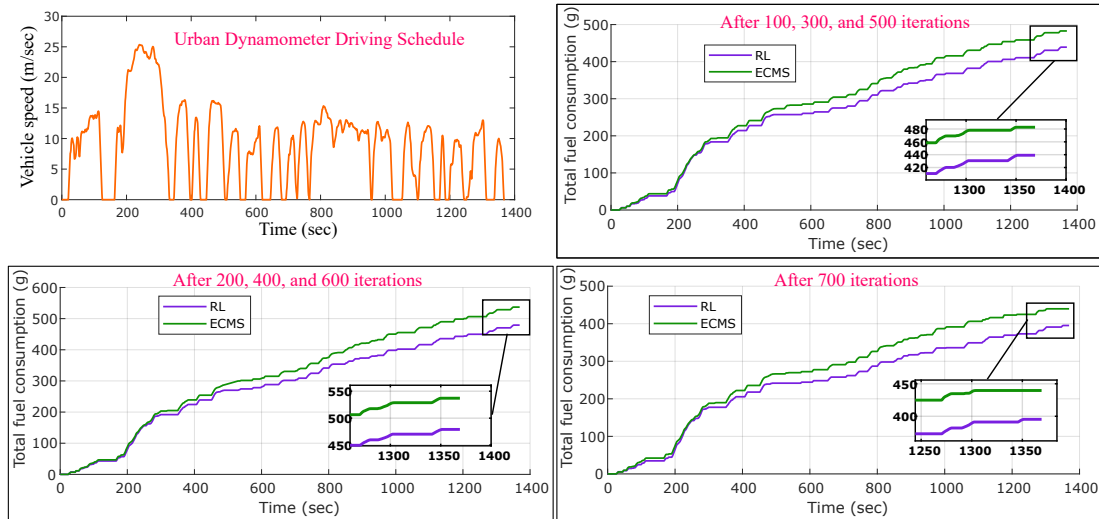


Figure 7.27: Effect of training iteration numbers on the agent’s fuel consumption performance compared to optimal ECMS for UDDS.

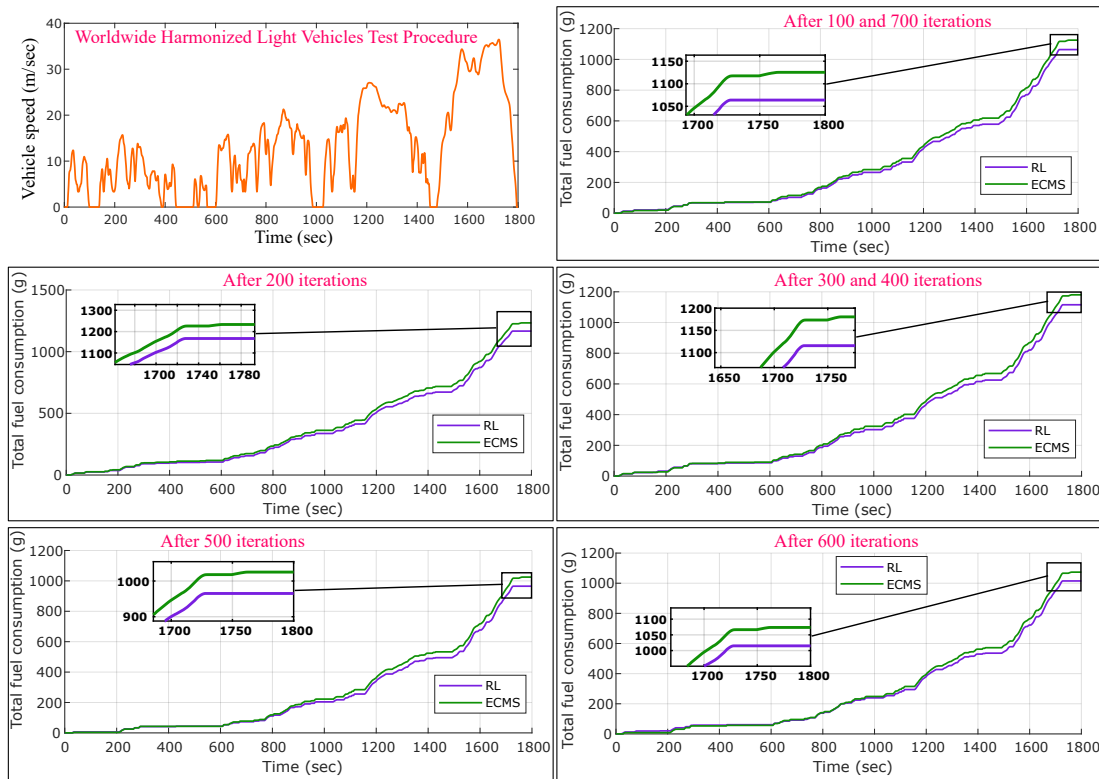


Figure 7.28: Effect of training iteration numbers on the agent’s fuel consumption performance compared to optimal ECMS for WLTP.

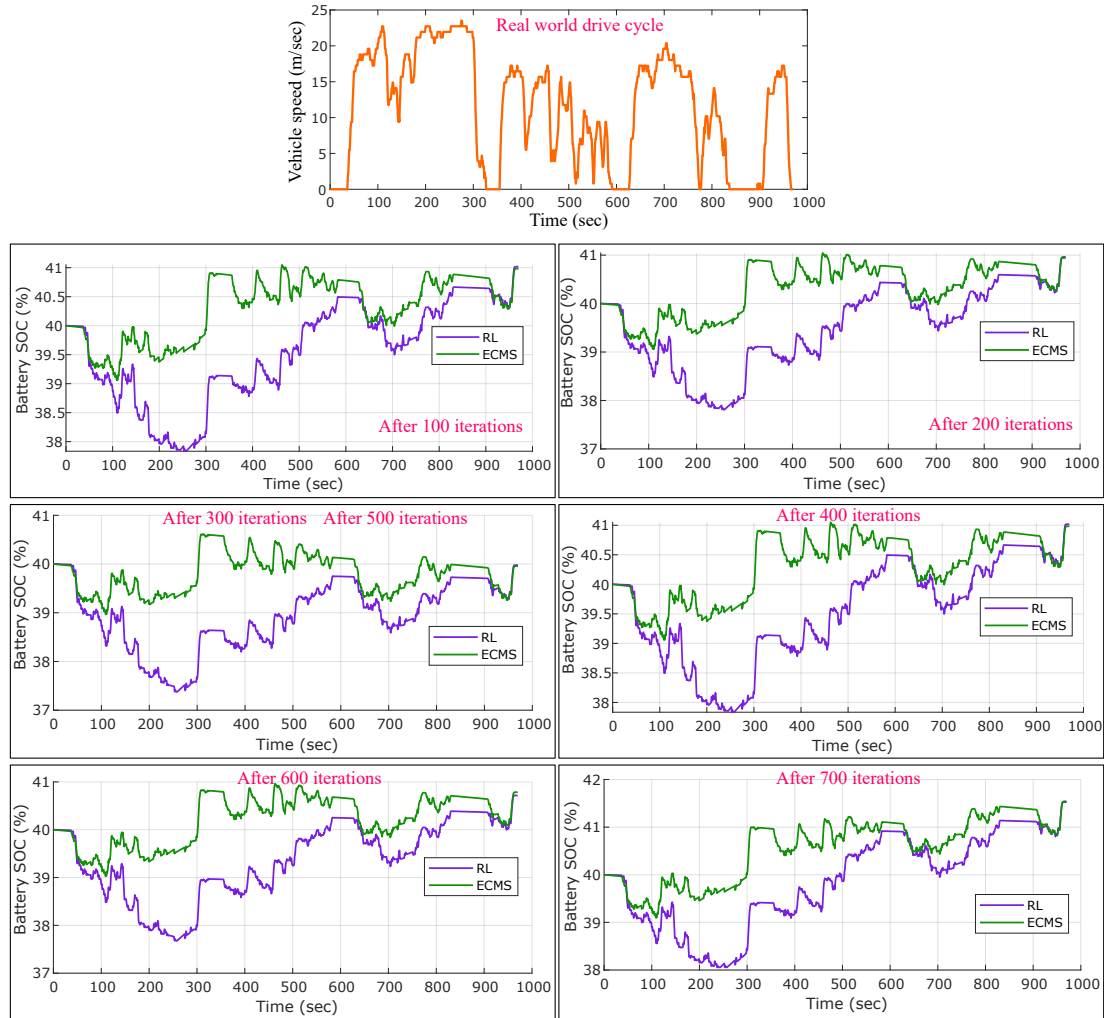


Figure 7.29: Effect of training iteration numbers on the agent’s charge sustenance capability. A real-world drive cycle is chosen here.

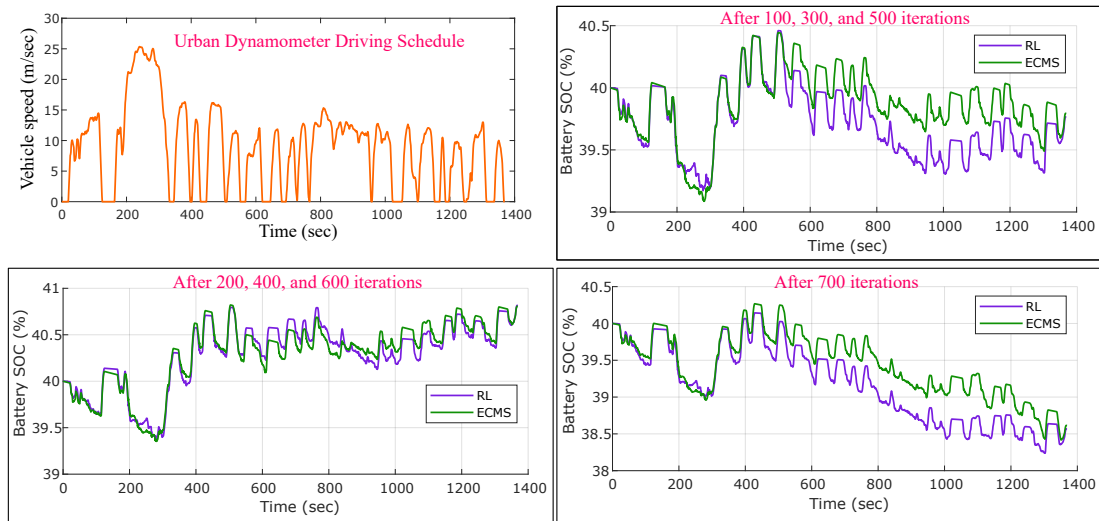


Figure 7.30: Effect of training iteration numbers on the agent’s charge sustenance capability. UDDS is chosen here.

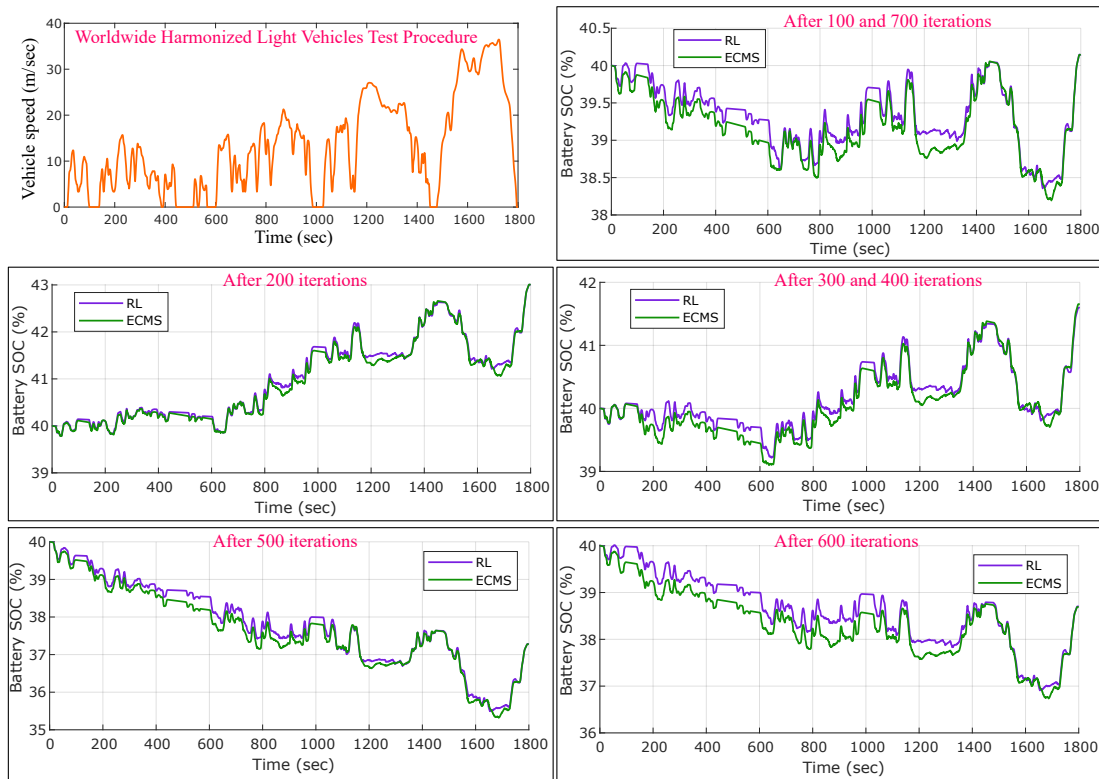


Figure 7.31: Effect of training iteration numbers on the agent’s charge sustenance capability. WLTP is chosen here.

## Chapter 8

# Advanced Modeling of Multi-mode Electrified Powertrain: Requirement of a Dynamic Coordinated Control to Manage the Mode-shift Events Efficiently



## 8.1 Introduction

An online simulation framework is developed in this chapter to evaluate the performance of a multi-mode electrified powertrain equipped with more than one power source. An electrically variable transmission with two planetary gear-set has been chosen as the centerpiece of the powertrain considering the versatility and prospects of such transmissions. A novel architecture topology of the aforementioned class of transmission is selected through rigorous screening process whose workflow is presented here with brevity. The article systematically delineates the steps for deriving dynamics associated with all the feasible operating modes facilitated by the selected topology. The dynamics associated with all the feasible mode-shift events are also heeded judiciously. One of the legitimate concern of multi-mode transmission is its proclivity to contribute discontinuity of power-flow downstream of the powertrain. Mode-shift events can be predominantly held responsible for engendering such discontinuity. Many scholars in literature have substantiated the advent of dynamic coordinated control as a technique for ameliorating such discontinuity. Hence, a system-level coordinated control is employed within the energy management system (equivalent consumption minimization strategy), which governs the mode schedule of the multi-mode powertrain in real-time simulation. Simulation results corroborate the effect of coordinated control on the equivalent consumption minimization strategy in generating optimal mode schedule.

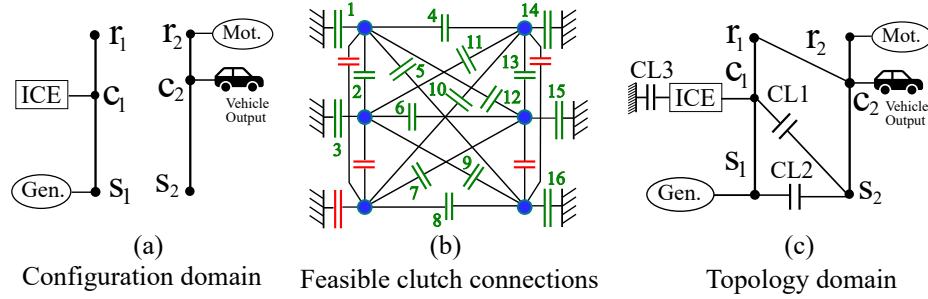


Figure 8.1: Example of a specific powertrain topology derived from a specific powertrain configuration.

## 8.2 Motivation

The second and third generations of THS with power-split configuration have been the most successful and ingenious ePTs throughout the last decade, as shown in Fig.1.2. Although many factors govern the success or defeat of any HEV in terms of sales figure, sticker value of fuel-economy, torque, and power rating of the HEV play a significant role in its success in the market. GM introduced the *Two-mode* hybrid system facilitating multiple electrified modes with two PG-sets and multiple clutches [6] in 2007. The *Two-mode* ePT architecture combines the advantages of both parallel and series architectures in a single unit, making it even more versatile than power-split [7]. However, the mode scheduling control becomes an added responsibility for the EMS whose control strategy needs to be even more complicated than that of a series or parallel or even power-split architecture. Moreover, the stark contrast between the success of the aforementioned two drivetrains, as shown in Fig.1.2, has ignited the search for advanced hybrid drivetrain configurations.

The search method’s purpose is two-folded, i.e., finding all the mathematically possible topology candidates and then screening out the best performing candidates. Fig.8.1.a and 8.1.c depict the conceptual difference between powertrain configuration

and powertrain topology, whereas Fig.8.1.b elucidates all the feasible clutch connections in a two PG-set configuration. Authors in [180] have shown how enormous the design space can transpire if only two PG-sets, each of which carrying two components, are allocated with three clutches among sixteen possible node connections. Liu and Peng (2010) [38] and authors in [39] presented a systematic and chronological screening methodology for selecting the best performing topology candidates from a large pool of mathematically possible topology candidates. Since the design space is remarkably large, an agile and an offline near-optimal control strategy called slope-weighted energy-based rapid control analysis (SERCA) has been appointed to evaluate the performance of the topology candidates [181] before screening out the best ones.

During the topology screening stage, it is not viable to assess the performance of even every feasible topology candidates through an online simulation considering the massiveness of the design space. But, the best-selected topology candidates can be modeled in an online simulation platform (OSP) to corroborate their claimed performance. From a practical point of view, any candidate chosen by any of the offline algorithms, must be validated with online mode scheduling control before subsequent validation stages such as SIL, HIL, and dyno test [4]. The works presented in [39,181] ends with an avenue for the development of a near-optimal control-based OSP to validate the fuel consumption value and drivability of the best candidates yielded through the automated topology screening process. Each of the best performing topology candidates with two PG-sets, three prime-movers, and three clutches will inevitably have multiple electrified modes, i.e., a few all-electric modes and a few hybrid-electric modes. A real-time near-global optimal mode scheduling control is indispensable for

validating the topology candidate’s performance in the OSP.

The necessity of a real-time mode scheduling control has been skyrocketing since the ascent in research on ePTs in the past two decades. Copious researchers have been concentrating on achieving swift and smooth ICE activation during mode transition [182–184], ameliorating the adverse effect of ICE’s ripple torque during ICE deactivation [177], attenuating transient drive-shaft vibration with motor torque compensation control during mode shifts [185], and improving drivability [186]. Apart from these studies, Lei et al. [187] proposed a coordinated control to expedite the simultaneous mode transition and clutch shifting process in a full HEV, whereas Tang et al. applied [188] a motor torque compensation control to alleviate the unpleasant vibration engendered due to frequent ICE start-stop. Despite their irrefutable contributions, the studies mentioned above are limited within parallel, series-parallel, and power-split configurations, but not for multi-mode topology. These studies focus on a detailed analysis of an individual mode shift event rather than the optimal mode schedule for the entire drive cycle.

Among a few works on mode scheduling control, Ahn and Cha (2008) [189] introduced the Pareto-frontier optimization for deciding optimal mode schedule, whereas Karbowski, Kwan, Kim, and Rousseau (2010) [190] employed simple rule-based control to optimize the mode schedule. Kim, Kim, Min, Hwang, and Kim (2011) [7] have only indicated the significance of clutch dynamics but have not included it while articulating an offline mode scheduling control for GM’s *Two-mode* hybrid system. Zhuang, Zhang, Yin, Peng, and Wang (2020) [191] have proposed an integrated real-time mode scheduling strategy by amalgamating a novel *normalized efficiency maximum strategy* with an optimal mode-shift map generation. However, the strict

constraint on feasible clutch engagement makes the EMS controller unsuitable for unfamiliar drive cycles. The ingenuity lies in the employment of Dijkstra’s algorithm for designing the shortest energy-path for indirect mode-shifts while considering ride comfort. However, the proposed method imposes a strict constraint, i.e., equality of both discs’ speed, for a feasible clutch engagement, making the EMS controller conservative and eventually making the EMS controller unsuitable for unfamiliar drive cycles. Moreover, the optimal mode-shift schedule is obtained by DP, making the overall method computationally expensive and time-consuming. Zhuang, Zhang, Li, Wang, and Yin (2017) [142] proffered a real-time mode scheduling control by conflating offline and online computation. However, the optimal mode-shift map must be computed for every drive cycle separately, making the entire method unrealistic for real-world drive cycles. Geng, Lou, Wang, and Zhang (2020) [192] proposed a composite implementation of DP and ECMS for mode scheduling control. However, it is not real-time implementable. Anselma, Hou, Roeleveld, Belingardi, and Emadi (2019) [193] employed an ANN, trained with near-optimal mode schedule from offline results, as an online implementable EMS controller. Buccoliero, Anselma, Bonab, Belingardi, and Emadi (2020) [194] proposed an improved version of SERCA, yielding near-global optimal mode scheduling control for multi-mode ePTs. However, no transient powertrain dynamics has been considered while articulating the mode scheduling control. The assessment of fuel economy and operational feasibility of a multi-mode ePT is incomplete without validating its transient performance during mode-shift events.

Looking at the scarcity of research on real-time mode scheduling control for multi-mode ePTs, the current endeavor proffering a near-global optimal and real-time mode

scheduling control for a novel ePT topology will seek to quench the research gap. The topology is yielded from the work presented in [39]. Compared to other feasible mode scheduling control in the literature, the current proposed real-time mode scheduling control encompasses the effect of transient powertrain dynamics, drivability improvement, feasibility of mode shifts, and near-global optimality. The original contributions of this article are enumerated as follows: 1) Dynamics of every mode of a novel ePT topology is modeled in detail with inertia-based approach instead of a steady-state approach. 2) Dynamics associated with every feasible mode-shift is separately modeled to illustrate the clutch-to-clutch shift dynamics. The mode-shift dynamics has a significant influence in generating the optimal mode-shift schedule. 3) Low-frequency coordinated control is proposed for the first time, to the best of the authors' knowledge, not only for the all-electric to hybrid-electric mode-shifts but also the hybrid-electric to hybrid-electric shifts and all-electric to all-electric shifts. The coordinated control plays a pivotal role in emulating a real physical ePT with an inertia-based dynamic ePT model. 4) Finally, an instantaneous optimization, influenced by the coordinated control, is proposed for the novel topology to derive a near-global optimal and feasible mode-shift schedule for any given drive cycle.

### 8.3 Automated Topology Selection Tool

The automated topology selection tool (ATST) is comprised of two major phases. First, the design space exploration phase where all the feasible topology candidates with a certain number of PG-sets, components (ICE and EMGs), and clutches are

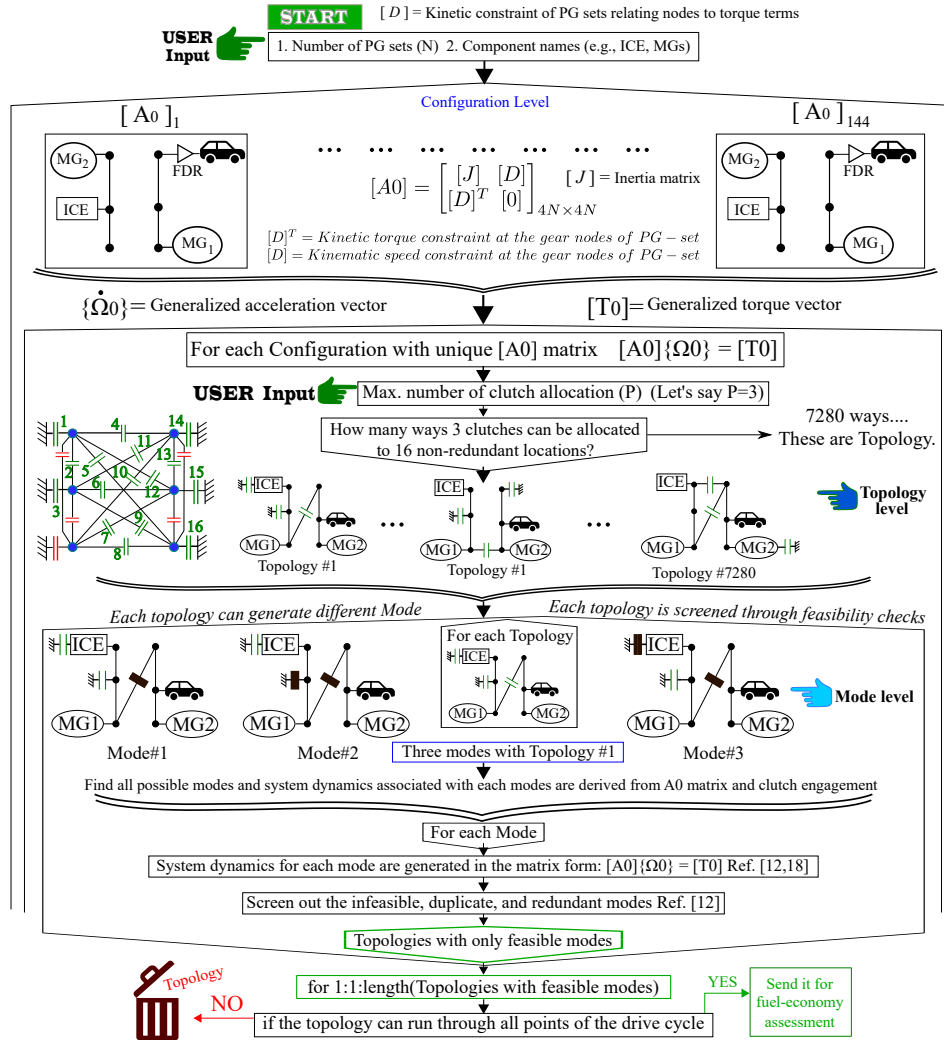


Figure 8.2: Automatic and exhaustive topology candidate screening process.

generated through permutation and combinations of both component and clutch allocations. Second, an exhaustive screening phase, which facilitates a two-stage screening, i.e., eliminating the topology candidates with infeasible, redundant, and duplicate modes and then elimination of the candidates with inferior performance. An offline control strategy named SERCA is employed to assess the performance of every feasible candidate.

### 8.3.1 Exhaustive Search of Topology Candidates

The design space exploration is a continuation of a previous work presented in [39] which was inspired by [195]. Since this article focuses directly on the performance of the topology candidates yielded from the ATST, both exploration and assessment phases are not discussed in detail. A comprehensive understanding of the whole exploration phase is depicted in Fig.8.2. Interested readers are referred to the article [39]. Tab. 8.1 tabulates the specifications of the powertrain components and vehicle components used in the automated topology screening process.

### 8.3.2 Screening of Topology Candidates

SERCA algorithm, which was initially proposed by Anselma et al. [181], is employed as the offline control strategy for assessing the performance of every feasible candidate. SERCA can yield very close to the global optimal performance given by DP in much lesser computational time. Due to its computational agility, SERCA has been preferred over DP for rapid assessment of fuel economy capability of the topology candidates with multiple PG-sets as well as conventional electrified architectures [196]. Fig.8.3 shows the outcome of a rapid performance assessment of more than 100,000 topology candidates. The performance metric is chosen as the average fuel consumption over three standard federal drive cycles, i.e., UDDS, highway cycle, and WLTC. Any red circle in Fig.8.3 represents the average fuel consumption of any given topology candidates. The only blue square represents a state-of-the-art ePT called single input electrically variable transmission (Si-EVT).



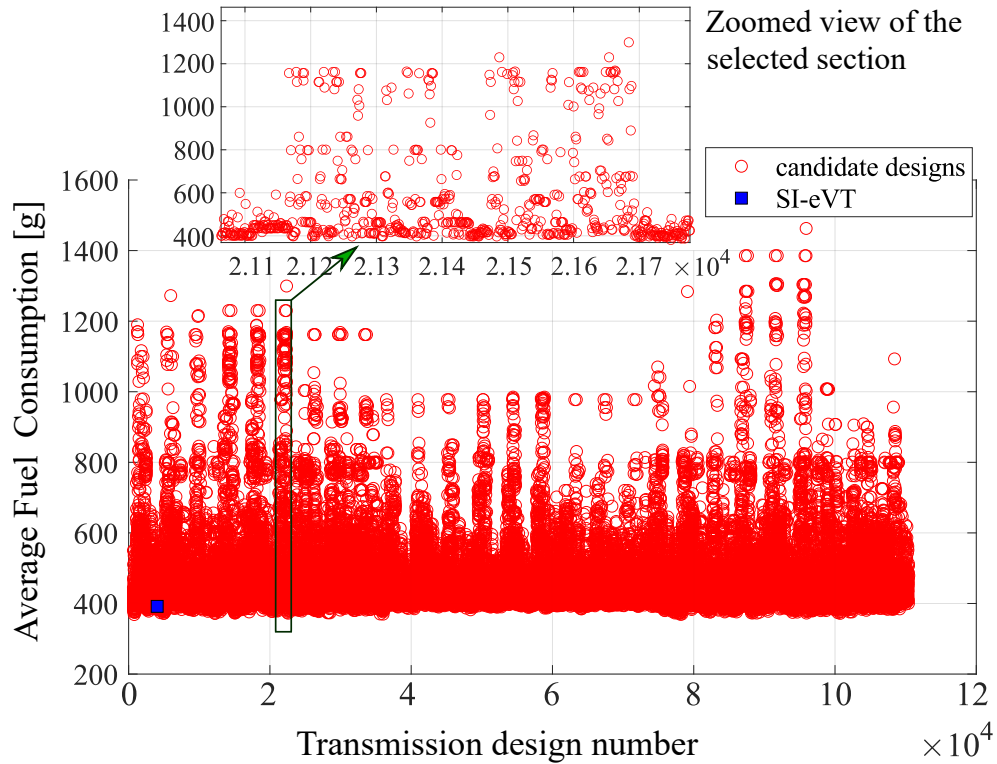


Figure 8.3: Outcome of a rapid assessment of fuel economy capability of more than 100,000 topology candidates.

## 8.4 Powertrain Modeling

### 8.4.1 Modeling of Dynamics for Different Modes

The multi-mode powertrain topology, as shown in Fig.8.4, has been selected from a pool of well-performing topology candidates among approximately 100,000 design candidates. Equipped with two PG-sets, one ICE, two EMGs, i.e., a traction motor and a generator, and three clutches, the topology can facilitate four distinct powertrain operating modes, i.e., two all-electric modes and two hybrid-electric modes. The clutch connections required for all the four modes are given in Tab.8.2, where  $u_{control}$  refers to the control variables corresponding to the modes. The ICE is connected to

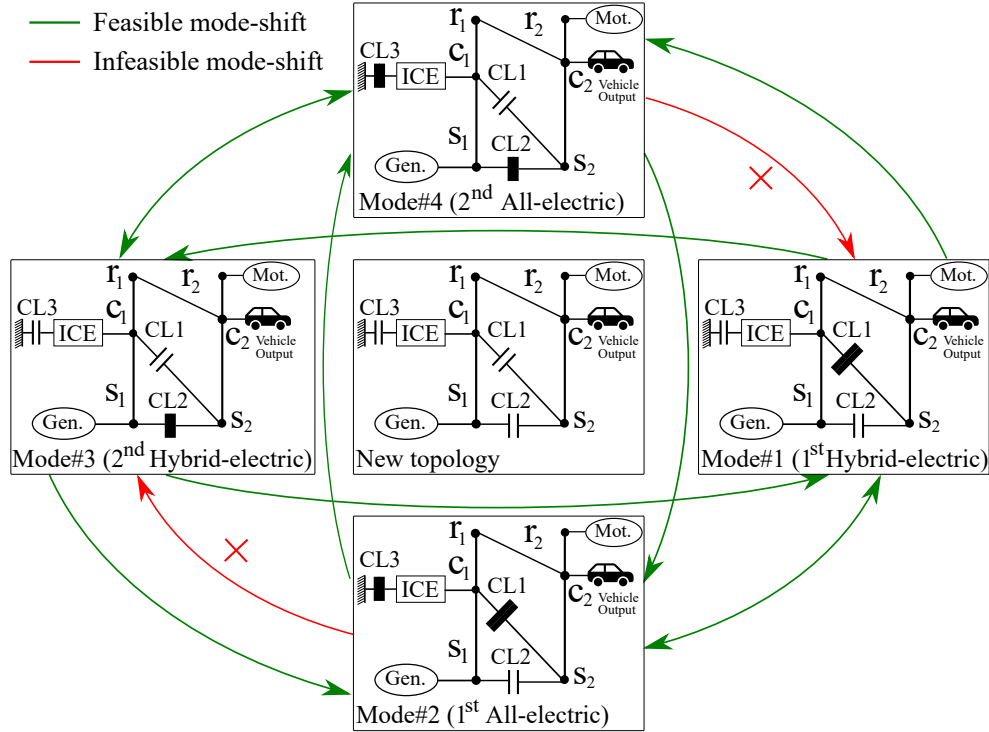


Figure 8.4: Mode-shifts among the four feasible modes.

the C1 through a OWC (CL#3), which prevents the ICE from rotating anti-clockwise when the generator applies propelling torque in all-electric modes. The generator and the traction motor are connected to S1 and R2, respectively. The final drive is connected to C2.

### Dynamics of Powertrain in Mode#1

In mode#1, the traction motor is connected to only R2, and the generator is connected to only S1. Hence, the dynamics associated with S1 and R2 can be written as Eq.8.4.1a. But, a similar equation will not be appropriate for C1, which is connected to both ICE and S2. Hence,  $\tau_{c1} \neq \tau_{ice} - J_{c1}\ddot{\theta}_{ice}$ . Instead, the dynamics associated with C1 is expressed with Eq.8.4.1e. Similarly, it will be incorrect to

write  $\tau_{c2} = \tau_{out} - J_{c2}\ddot{\theta}_{out}$  since C2 is connected to both transmission output and R1. The dynamics associated with output is expressed through Eq.8.4.1d. The dynamics associated with R1, C2, and S2 can be expressed through abiding by the torque relationship within PG-set#1 and PG-set#2 respectively, as given by Eq.8.4.1b and Eq.8.4.1c respectively. Finally, dynamics of the output can be obtained by substituting Eq.8.4.1b and the first part of Eq.8.4.1c in Eq.8.4.1d. Similarly, the ICE dynamics can be obtained by substituting the second part of Eq.8.4.1c in Eq.8.4.1e.

The inertia-based dynamic relationship among different components of the powertrain (components of PG-set, ICE, generator, motor, and transmission output) in mode#1 can be expressed in matrix notation as given in Eq.8.4.2.

$$\tau_{s1} = \tau_{gen} - J_{s1}\ddot{\theta}_{gen} ; \tau_{r2} = \tau_{mot} - J_{r2}\ddot{\theta}_{mot} \quad (8.4.1a)$$

$$\tau_{r1} = \beta_1\tau_{s1} = \beta_1(\tau_{gen} - J_{s1}\ddot{\theta}_{gen}) \quad (8.4.1b)$$

$$\tau_{c2} = -\left(\frac{1 + \beta_2}{\beta_2}\right)(\tau_{r2}) ; \tau_{s2} = \frac{\tau_{r2}}{\beta_2} \quad (8.4.1c)$$

$$\tau_{out} = \tau_{r1} + \tau_{c2} + J_{out}\ddot{\theta}_{out} \quad (8.4.1d)$$

$$\tau_{c1} = \tau_{ice} - J_{c1}\ddot{\theta}_{ice} - \tau_{s2} ; \ddot{\theta}_{ice} = \ddot{\theta}_{S2} \quad (8.4.1e)$$

$$\begin{bmatrix} \tau_A \\ \tau_B \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{c2} & 0 & -\beta_1 J_{s1} & \left(\frac{1+\beta_2}{\beta_2}\right)J_{r2} & 0 \\ 0 & J_{c1} & (1 + \beta_1)J_{s1} & -\frac{1}{\beta_2}J_{r2} & 0 \\ (1 + \beta_2) & -1 & 0 & -\beta_2 & 0 \\ \beta_1 & -(1 + \beta_1) & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_{out} \\ \ddot{\theta}_{ice} \\ \ddot{\theta}_{gen} \\ \ddot{\theta}_{mot} \\ \ddot{\theta}_{S2} \end{bmatrix} \quad (8.4.2)$$

, where  $\tau_B = \tau_{ice} + (1 + \beta_1)\tau_{gen} - \frac{1}{\beta_2}\tau_{mot}$ ,  $\tau_A = \tau_{out} - \beta_1\tau_{gen} + \left(\frac{1+\beta_2}{\beta_2}\right)\tau_{mot}$ ,  $J_{c2} =$

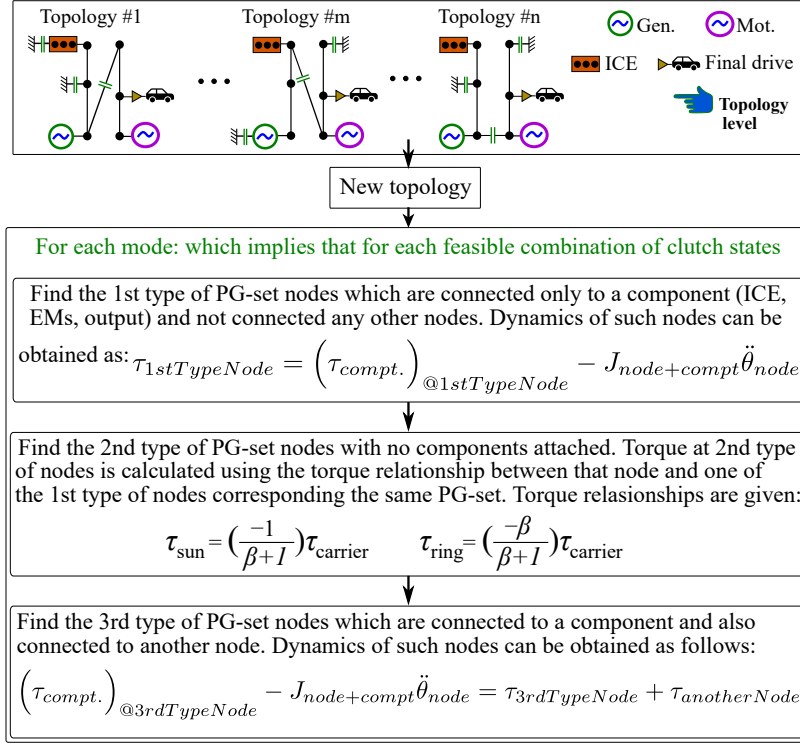


Figure 8.5: Systematic method of obtaining dynamic relationship among components and nodes of an ePT with multiple PG-sets.

$J_{C2} + J_{out} + J_{R1}$ ,  $J_{s1} = J_{S1} + J_{gen}$ ,  $J_{r2} = J_{R2} + J_{mot}$ ,  $J_{c1} = J_{C1} + J_{ice}$ .  $J_C$ ,  $J_S$ , and  $J_R$  are the MMIs of the carrier, sun gear, and the ring gear of PG-set, respectively.  $J_{out}$  is the effective combined-MMI of vehicle, wheels, and final drive measured at the transmission output.  $J_{ice}$ ,  $J_{gen}$ , and  $J_{mot}$  are the MMIs of ICE, generator, and traction motor, respectively.  $\beta_1 = \frac{N_{R1}}{N_{S1}}$  and  $\beta_2 = \frac{N_{R2}}{N_{S2}}$ , where  $N_R$  and  $N_S$  are the number of teeth of the ring gear and sun gear, respectively. Eq.8.4.1 has been expressed in more compact notation with Eq.8.4.2.

## Dynamics of Powertrain in Mode#2

The powertrain dynamics in mode#2 is obtained through the same concept used for mode#1. The systematic method of obtaining the powertrain dynamics of an operating mode of multi-mode e-CVT transaxle is streamlined and elucidated in Fig.8.5. It is noteworthy that both the method illustrated in Fig.8.5 and the methods prescribed in [195], and [197] yield precisely the same powertrain dynamics in every mode. While the theoretical method proposed in [38] and [195] has been widely acclaimed by many scholars, authors in [197] validated the theoretical method through experiment. Hence, the method illustrated in Fig.8.5 does not need any experimental validation separately. Since the CL#3 is engaged, C1 is assumed to be rigidly connected to the ground. As per the systematic approach elucidated in Fig.8.5, R2 and S1 satisfy the conditions of the first type of PG-set nodes. Hence, the dynamics corresponding to R2 and S1 can be given by Eq.8.4.3a. R1 and C2 fall under the category of the second type and the third type of PG-set nodes, respectively, and dynamics of R1 can be given by Eq.8.4.3b. Since the output is connected to both C2 and R1, dynamics associated with output can be given by first part of Eq.8.4.3d.

$$\tau_{s1} = \tau_{gen} - J_{s1}\ddot{\theta}_{gen} ; \tau_{r2} = \tau_{mot} - J_{r2}\ddot{\theta}_{mot} \quad (8.4.3a)$$

$$\tau_{r1} = \beta_1\tau_{s1} = \beta_1(\tau_{gen} - J_{s1}\ddot{\theta}_{gen}) \quad (8.4.3b)$$

$$\tau_{c2} = -\left(\frac{1 + \beta_2}{\beta_2}\right)(\tau_{mot} - J_{r2}\ddot{\theta}_{mot}) \quad (8.4.3c)$$

$$\tau_{out} = \tau_{r1} + \tau_{c2} + J_{out}\ddot{\theta}_{out} ; \ddot{\theta}_{S2} = 0 \quad (8.4.3d)$$

Inertia-based dynamic relationship in the mode#2 (1<sup>st</sup> all-electric mode) can be expressed in more compact form:

$$\begin{bmatrix} \tau_A \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{c2} & -\beta_1 J_{s1} & (\frac{1+\beta_2}{\beta_2}) J_{r2} & 0 \\ (1 + \beta_2) & 0 & -\beta_2 & 0 \\ \beta_1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_{out} \\ \ddot{\theta}_{gen} \\ \ddot{\theta}_{mot} \\ \ddot{\theta}_{S2} \end{bmatrix} \quad (8.4.4)$$

### Dynamics of Powertrain in Mode#3

The powertrain dynamics of the mode#3 is obtained as Eq.8.4.5 by following the generic method depicted in Fig.8.5.

$$\tau_{c1} = \tau_{ice} - J_{c1}\ddot{\theta}_{ice} ; \tau_{r2} = \tau_{mot} - J_{r2}\ddot{\theta}_{mot} \quad (8.4.5a)$$

$$\tau_{r1} = -\left(\frac{\beta_1}{1 + \beta_1}\right)(\tau_{ice} - J_{c1}\ddot{\theta}_{ice}) \quad (8.4.5b)$$

$$\tau_{c2} = -\left(\frac{1 + \beta_2}{\beta_2}\right)(\tau_{mot} - J_{r2}\ddot{\theta}_{mot}) \quad (8.4.5c)$$

$$\tau_{out} = \tau_{r1} + \tau_{c2} + J_{out}\ddot{\theta}_{out} \quad (8.4.5d)$$

$$\tau_{s1} = -\left(\frac{\beta_1}{1 + \beta_1}\right)(\tau_{ice} - J_{c1}\ddot{\theta}_{ice}) \quad (8.4.5e)$$

$$\tau_{s2} = \frac{1}{\beta_2}\tau_{r2} = \frac{1}{\beta_2}(\tau_{mot} - J_{r2}\ddot{\theta}_{mot}) \quad (8.4.5f)$$

$$\tau_{gen} = \tau_{s1} + \tau_{s2} + J_{s1}\ddot{\theta}_{gen} ; \ddot{\theta}_{gen} = \ddot{\theta}_{S2} \quad (8.4.5g)$$

Inertia-based system dynamics for mode#3 can be expressed in more compact notation as follows:

$$\begin{bmatrix} \tau_C \\ \tau_D \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{c2} & (\frac{\beta_1}{1+\beta_1})J_{c1} & 0 & (\frac{1+\beta_2}{\beta_2})J_{r2} & 0 \\ 0 & (\frac{1}{1+\beta_1})J_{c1} & (J_{s1} + J_{S2}) & -\frac{1}{\beta_2}J_{r2} & 0 \\ (1 + \beta_2) & 0 & -1 & -\beta_2 & 0 \\ \beta_1 & -(1 + \beta_1) & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_{out} \\ \ddot{\theta}_{ice} \\ \ddot{\theta}_{gen} \\ \ddot{\theta}_{mot} \\ \ddot{\theta}_{S2} \end{bmatrix} \quad (8.4.6)$$

, where  $\tau_C = \tau_{out} + (\frac{\beta_1}{1+\beta_1})\tau_{ice} + (\frac{1+\beta_2}{\beta_2})\tau_{mot}$  and  $\tau_D = (\frac{1}{1+\beta_1})\tau_{ice} + \tau_{gen} - \frac{1}{\beta_2}\tau_{mot}$ .

#### Dynamics of Powertrain in Mode#4

The powertrain dynamics of mode#4 is obtained as Eq.8.4.7 by following the generic method depicted in Fig.8.5.

$$\tau_{r2} = \tau_{mot} - J_{r2}\ddot{\theta}_{mot} \quad (8.4.7a)$$

$$\tau_{s1} = \tau_{gen} - J_{s1}\ddot{\theta}_{gen} - \tau_{s2} ; \tau_{r1} = \beta_1\tau_{s1} \quad (8.4.7b)$$

$$\tau_{s2} = \frac{1}{\beta_2}\tau_{r2} = \frac{1}{\beta_2}(\tau_{mot} - J_{r2}\ddot{\theta}_{mot}) \quad (8.4.7c)$$

$$\tau_{c2} = -(\frac{1 + \beta_2}{\beta_2})(\tau_{mot} - J_{r2}\ddot{\theta}_{mot}) \quad (8.4.7d)$$

$$\tau_{out} = \tau_{r1} + \tau_{c2} + J_{out}\ddot{\theta}_{out} ; \ddot{\theta}_{S2} = \ddot{\theta}_{s1} \quad (8.4.7e)$$

Inertia-based system dynamics for mode#4 can be expressed in matrix notation as follows:

$$\begin{bmatrix} \tau_E \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_{c2} & -\beta_1(J_{s1} + J_{S2}) & (\frac{1+\beta_1+\beta_2}{\beta_2})J_{r2} & 0 \\ (1 + \beta_2) & 0 & -\beta_2 & -1 \\ \beta_1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_{out} \\ \ddot{\theta}_{gen} \\ \ddot{\theta}_{mot} \\ \ddot{\theta}_{S2} \end{bmatrix} \quad (8.4.8)$$

, where  $\tau_E = \tau_{out} - \beta_1\tau_{gen} + (\frac{1+\beta_1+\beta_2}{\beta_2})\tau_{mot}$ . It is noteworthy to mention that special attention has been paid for modeling the dynamics of the node S2 in all the four modes. Apparently, the node S2 might seem to possess normal importance as the other nodes do, but the reason for paying heed to dynamics of S2 will soon be unveiled.

## 8.5 Clutch-based Coordinated Control during Mode-shifts

Multi-mode ePT can be contrived with not only multiple hybrid-electric modes but multiple all-electric modes also. Alike automatic transmission, clutch-based mode-shift operation has been soaring as the predominant field for encompassing many electrified modes into a single ePT topology. The clutch operation associated with a mode-shift event interrupts the desired flow of torque to the transmission output due to the discontinuous and highly nonlinear aspects of its torque dynamics. Traditionally, the obligation of DCCS is two-folded. The first one is to compensate the deficiency of torque, which is caused due to the slow response of ICE during ICE



start-up, in the output with fast responding EMG [185]. The second one is to eradicate the effect of ICE ripple or resistance torque in the transmission output through the agility of EMG's torque [198].

Throughout this decade, literature has been inundated with a copious number of research works presenting the ingenuity of DCCS in eradicating the inconsistent power flow at the output during mode-shifts. DCCS has been predominantly applied during the transition from all-electric mode to hybrid-electric mode [184], [187], [182]. Zhao and Tang (2019) [177] applied DCCS during hybrid-electric to all-electric mode-shift, i.e., during ICE-stopping phase. However, there is a scarcity of research work pertaining to the application of DCCS during all-electric to all-electric mode-shifts, hybrid-electric to hybrid-electric mode-shifts. Friction clutch based mode-shifts have been considered as a complicated controlled operation due to associated nonlinear dynamics in the slipping phase [183,199]. Although the effect of ICE ripple torque would be absent during either all-electric to all-electric or hybrid-electric to hybrid-electric mode-shifts, the ubiquitous effect of clutch slipping stage would tend to deteriorate the desired torque dynamics at the transmission output. Looking at the conspicuous qualities of DCCS, it is highly recommended to employ a DCCS in all kind of feasible mode-shifts, i.e., all-electric to all-electric shifts, all-electric to hybrid-electric shifts, hybrid-electric to all-electric shifts, and hybrid-electric to hybrid-electric shifts.

One of the cardinal contributions of this article is to demonstrate the effect of DCCS on the performance of EMS in terms of mode scheduling control. In this article, the EMS imbibes a rudimentary version of DCCS while controlling the mode schedule.

Fig.8.4 differentiates between two categories of mode-shifts. The feasibility of a mode-shift is dictated by the number of simultaneous clutch engagement and disengagement.

Tab.8.3 dictates that a feasible mode-shift event can have only up to two simultaneous clutch operations. That is why 4-to-1 and 2-to-3 mode-shifts are not feasible since they require three clutch operations simultaneously. Zhuang et al. [191] have labeled infeasible mode-shifts as indirect ones, and their work is focused on finding the optimal energy efficient path for executing the indirect mode-shifts.

It is noteworthy to mention that this work is not free from assumptions. Since clutch-to-clutch shift dynamics is a complex nonlinear phenomena and since it is relatively new in ePT application, a few assumptions have been taken into consideration while incorporating the effect of clutch-to-clutch shift into ECMS-based real-time optimal EMS.

### 8.5.1 Clutch-based Coordinated control for 1-to-3 Mode-shift

While operating in mode#1, if the EMS controller decides to shift to mode#3, a clutch-to-clutch shift takes place. CL#1 and CL#2 will be the off-going clutch and incoming clutch, respectively, during the 1-to-3 mode-shift. It is assumed in this article that clutch-to-clutch shifts happen gradually rather than instantaneously. Fig.8.6 shows a typical situation that can take place during the 1-to-3 mode-shift.

When transmission output speed, as shown in Fig.8.6, at the C2 ( $\omega_{out}$  or  $\omega_{C2}$ ) is lesser than the speed of ICE ( $\omega_{ice}$ ), generator speed ( $\omega_{gen}$ ) is greater than  $\omega_{ice}$  due to kinematics of PG-set. Since the clutch between the C1 and S2 is fully engaged,  $\omega_{S2}$  is equal to  $\omega_{ice}$  in mode#1. Hence,  $\omega_{S1} > \omega_{S2}$  in this mode for the aforementioned situation. CL#1 starts off-going, and CL#2 starts oncoming a single time-step after the ECMS-based EMS commands for the 1-to-3 mode shift. It is assumed that the torque transfer through the off-going clutch plummets to zero immediately after the

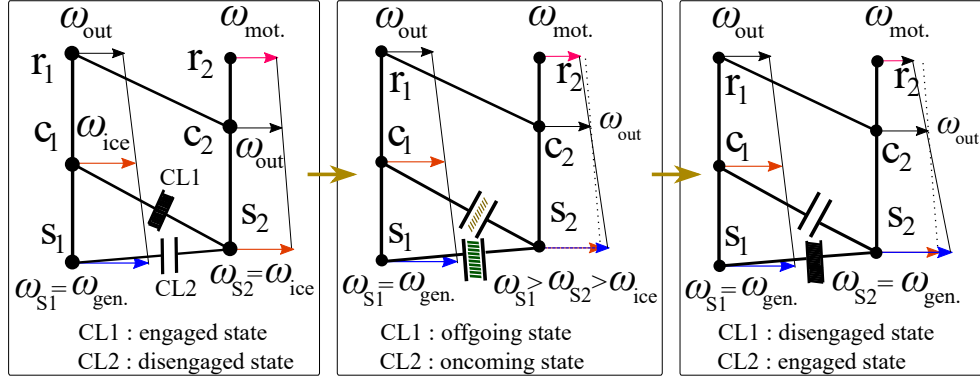


Figure 8.6: Different stages of clutch-to-clutch shift from mode#1 to mode#3 in a case when  $\omega_{ice}$  is greater than  $\omega_{out}$ .

disengagement happens. However, the torque transfer for the oncoming clutch continues until both the discs of the oncoming clutch achieve a unanimous rotational speed. Hence, the effect of  $\tau_{CL1}$  diminishes immediately after  $\tau_{CL2}$  activates. Moreover, the effect of  $\tau_{CL2}$  persists until  $\omega_{S2}$  equalizes to  $\omega_{S1}$ . It is very important to deduce the equations that describe the powertrain dynamics during the 1-to-3 mode shift.

At the very beginning of the 1-to-3 mode-shift, it is assumed that CL#1 goes through the “off-going” phase and still the torque transfer between S2 and C1 takes place. It is also assumed that CL#2 enters the very early phase of “oncoming” during this beginning. During this time, the powertrain dynamics can be given by the equations shown in the left-bottom of Fig.8.7.

As mentioned earlier, the clutch-to-clutch shift has two distinct phases, i.e., load transfer or torque transfer phase and inertia transfer or speed synchronization phase [200]. To keep the complexity within a reasonable limit, it is also assumed that torque transfer happens instantaneously from off-going clutch to oncoming clutch. The inertia transfer phase is modeled to be gradually happening. As shown in Fig.8.7, CL#1 mainly transmits  $\tau_{CL1} = \left(\frac{\tau_{r2}}{\beta_2}\right)$  amount of torque before its disengagement. A single time-step after the command of mode-shift, the responsibility of transmitting

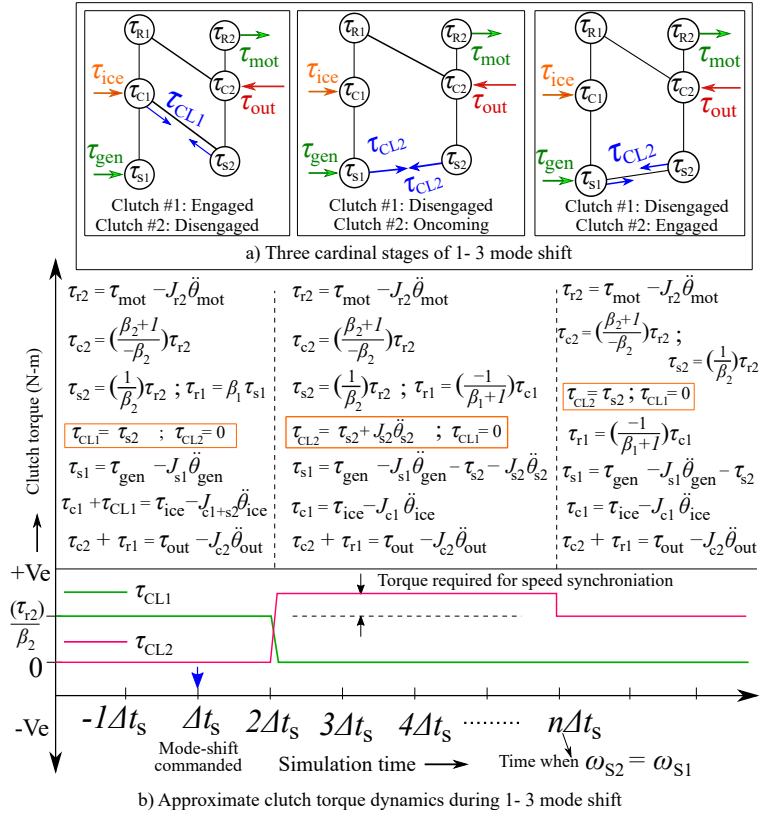


Figure 8.7: Three cardinal stages of torque transfer during mode-shift from mode#1 to mode#3.

( $\frac{\tau_{r2}}{\beta_2}$ ) amount of torque is delegated to CL#2 since CL#2 enters the oncoming phase. This delegation happens instantaneously, as shown in Fig.8.7. Now, CL#2 not only has to transmit the torque delegated to itself from CL#1, but CL#2 also has to transmit an extra amount of torque that is required for speed synchronization between  $\omega_{S1}$  and  $\omega_{S2}$ .

The powertrain dynamics during the oncoming phase of CL#2 is given by Eq.8.5.1.

$$\tau_{r2} = \tau_{mot} - J_{r2}\ddot{\theta}_{mot} ; \tau_{s2} = \frac{\tau_{r2}}{\beta_2} \quad (8.5.1a)$$

$$\tau_{c2} = -\left(\frac{1 + \beta_2}{\beta_2}\right)\tau_{r2} ; \tau_{r1} = -\left(\frac{1}{\beta_1}\right)\tau_{c1} \quad (8.5.1b)$$

$$\tau_{CL2} = \tau_{S2} - J_{S2}\ddot{\theta}_{S2} ; \tau_{CL1} = 0 \quad (8.5.1c)$$

$$\tau_{s1} = \tau_{gen} - J_{s1}\ddot{\theta}_{gen} - \tau_{S2} - J_{S2}\ddot{\theta}_{S2} \quad (8.5.1d)$$

$$\tau_{out} = \tau_{r1} + \tau_{c2} + J_{out}\ddot{\theta}_{out} \quad (8.5.1e)$$

$$\tau_{c1} = \tau_{ice} - J_{c1}\ddot{\theta}_{ice} \quad (8.5.1f)$$

Finally, the powertrain dynamics after the completion of speed synchronization is depicted in the bottom-right of Fig.8.7.

### 8.5.2 Clutch-based Coordinated control for 3-to-1 Mode-shift

As depicted in Fig.8.8.b, 3-to-1 mode-shift is also facilitated by a clutch-to-clutch shift between CL#2 and CL#1 as done in 1-to-3 mode-shift with the disengagement and engagement of clutches being done in reverse order.

As promised earlier, the importance of augmenting dynamics of S2 node can be highlighted while powertrain dynamics was derived for every mode in the previous section.

As depicted in Fig.8.8.a, Fig.8.7.a, and Fig.8.6, node S2 is the common node between CL#1 and CL#2 connection. This is why the dynamics of the S2 node plays a crucial role during the mode-shift events in this topology candidate.

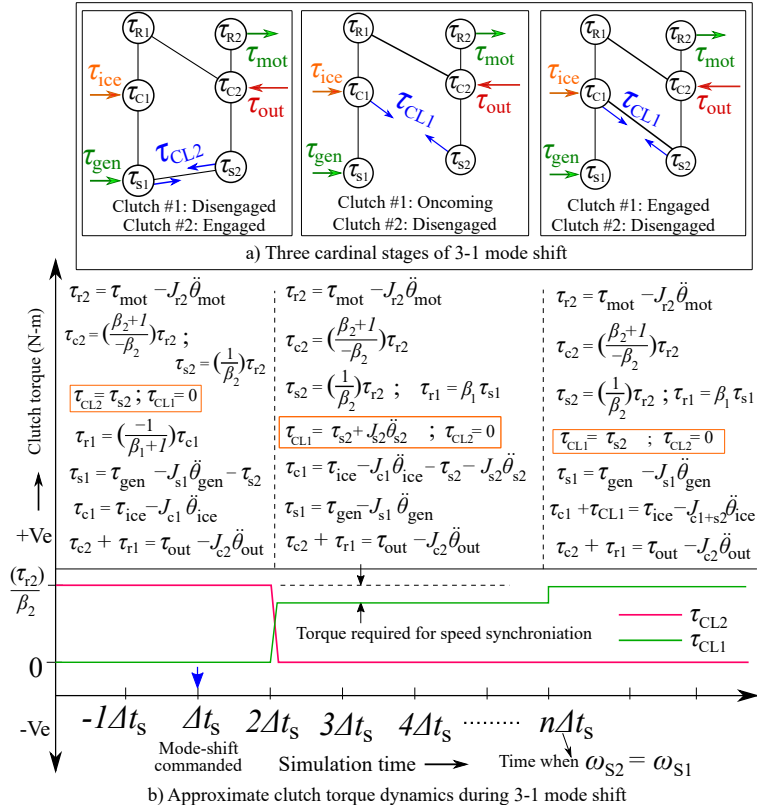


Figure 8.8: Three cardinal stages of torque transfer during mode-shift from mode#3 to mode#1.

### 8.5.3 Clutch-based Coordinated control for 2-to-4 Mode-shift

2-to-4 mode-shift represents an all-electric to all-electric mode-shift event. While normally operating in mode#2, CL#1 is fully engaged, i.e., S2 is not rotating at all like the C1 node. From the free body diagram (FBD) of S2 during mode#2, as shown in Fig.8.9.e, dynamics of the S2 node can be expressed with Eq.8.5.2.

$$\tau_{S2} - J_{S2}\ddot{\theta}_{S2} + \tau_{CL1} = 0 \quad (8.5.2)$$

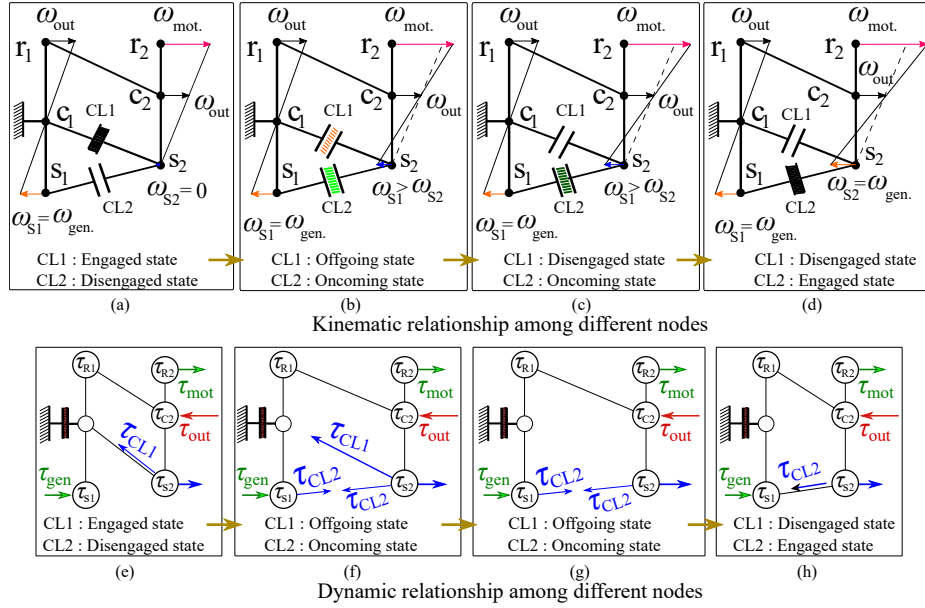


Figure 8.9: Kinematic and dynamic relationship among nodes of PG-sets during 2-to-4 mode-shift.

Zero acceleration of the S2 node during mode#2 implies that  $\tau_{CL1}$  is equal to  $-\tau_{s2}$  or  $-\frac{1}{\beta_2}\tau_{r2}$ .  $\tau_{CL1}$  provides the reaction torque to the S2 node so that the traction motor can propel the vehicle. Fig.8.9.b and Fig.8.9.f refer to the phase when “off-going” of CL#1 and “oncoming” of CL#2 commence simultaneously. As mentioned earlier, the off-going state of the disengaging clutch is assumed to be immediately perishing after the oncoming state of the engaging clutch starts. This is shown in Fig.8.9.c and Fig.8.9.g, where the oncoming state persists longer to enable the speed-synchronization gradually.

As shown in Fig.8.9.e and left side of Fig.8.10, the CL#1 primarily transmits  $\tau_{CL1} = -(\frac{\tau_{r2}}{\beta_2})$  amount of torque before it’s disengagement. A single time-step after the command of mode-shift, the responsibility of transmitting  $-(\frac{\tau_{r2}}{\beta_2})$  amount of torque is delegated to CL#2 since CL#2 enters the oncoming phase. This delegation happens

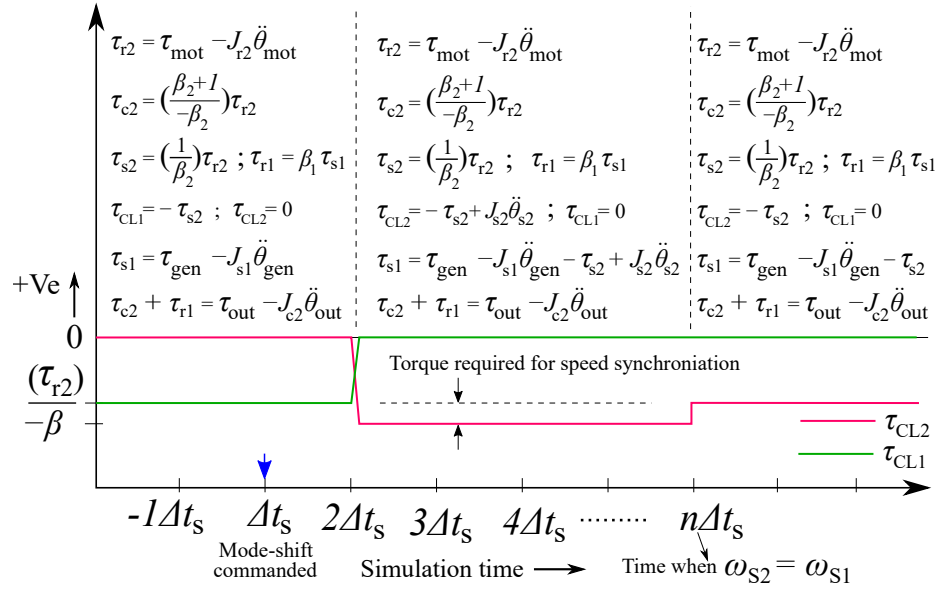


Figure 8.10: Approximate clutch torque variation of CL#1 and CL#2 during 2-to-4 mode-shift.

instantaneously, as shown in Fig.8.10. Now, CL#2 has to transmit the torque delegated to itself from CL#1, but CL#2 also has to transmit the extra amount of torque required for speed synchronization between  $\omega_{S1}$  and  $\omega_{S2}$ . Hence, the torque transmitted by CL#2 during the speed-synchronization phase is  $\tau_{CL2} = -\frac{1}{\beta_2}\tau_{r2} + J_{S2}\ddot{\theta}_{s2}$ . It can be observed in Eq.8.5.3 that there is an extra term  $\beta_1 J_{S2}\ddot{\theta}_{s2}$  in the output equation during the speed synchronization phase. This extra torque should be coordinated by the traction motor. Coordinated control is highly necessary during the speed synchronization phase so that the torque demand at the transmission output is properly satisfied at every time-steps during this phase.

$$\begin{aligned}
 (\tau_{out})_{during} &= J_{out}\ddot{\theta}_{out} + \beta_1(\tau_{gen} - J_{s1}\ddot{\theta}_{gen}) \\
 &\quad + \beta_1 J_{S2}\ddot{\theta}_{S2} - \left(\frac{1 + \beta_2 + \beta_1}{\beta_2}\right)(\tau_{r2})
 \end{aligned} \tag{8.5.3}$$



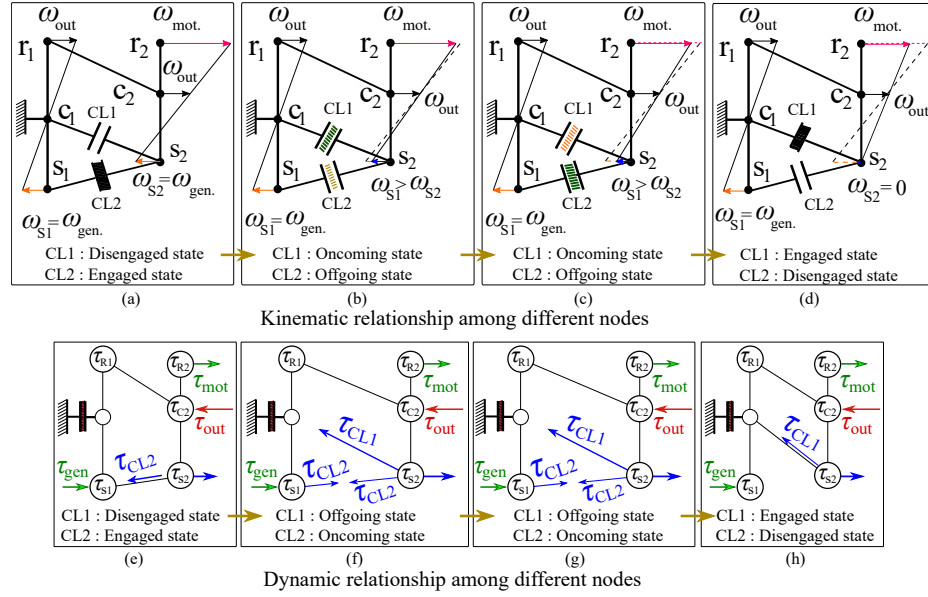


Figure 8.11: Kinematic and dynamic relationship among nodes of PG-sets during 4-to-2 mode-shift.

### 8.5.4 Clutch-based Coordinated control for 4-to-2 Mode-shift

During the 4-to-2 mode-shift,  $\omega_{S2}$  reduces from  $\omega_{S1}$  to zero value. CL#2 and CL#1 are the disengaging and engaging clutch, respectively. While Fig.8.11 depicts the kinematic and dynamic relationship among the nodes corresponding to different stages of mode-shift, Fig.8.12 elucidates the approximate variation in clutch (CL#1 and CL#2) torques during 4-to-2 mode-shift.

$$\begin{aligned}
 (\tau_{out})_{during} = & J_{out}\ddot{\theta}_{out} - \left(\frac{1 + \beta_2}{\beta_2}\right)(\tau_{mot} - J_{r2}\ddot{\theta}_{mot}) \\
 & + \beta_1(\tau_{gen} - J_{s1}\ddot{\theta}_{gen}) - (\beta_2 + 1)J_{S2}\ddot{\theta}_{S2}
 \end{aligned} \tag{8.5.4}$$

The coordinated control, as expressed in Eq.8.5.4, is advocated by traction motor, which is the primary torque controlling prime-mover in both the fully electric modes.

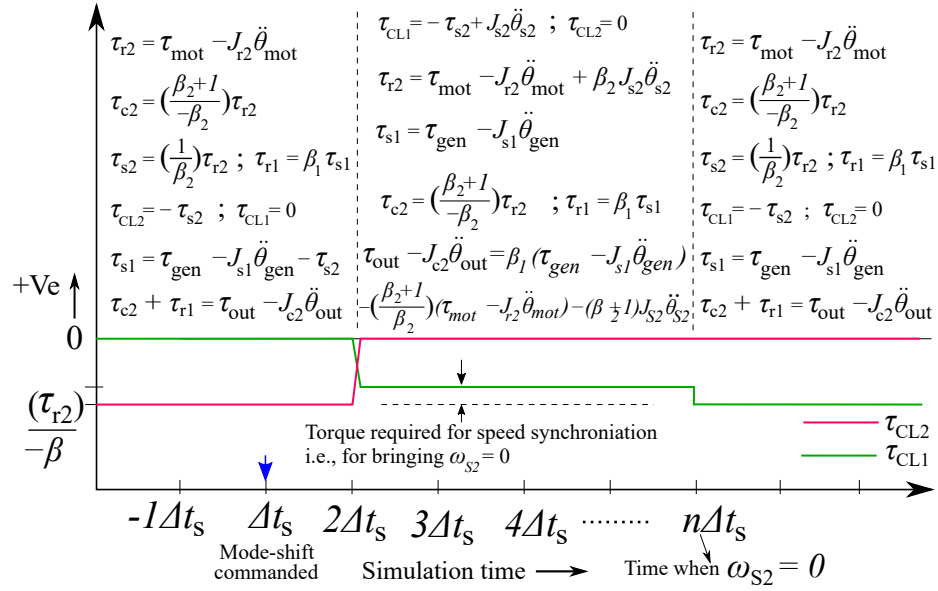


Figure 8.12: Approximate clutch torque variation of CL#1 and CL#2 during 4-to-2 mode-shift.

### 8.5.5 Coordinated control for All-electric to Hybrid-electric Mode-shift

The OWC allows the ICE to rotate only in the clockwise (conventional direction). Technically, while at the stand-still condition, the ICE can be cranked only clockwise up to ICE firing speed if generator applies positive torque. But, the ICE cannot be rotated anti-clockwise even if generator applies negative torque and, hence, the ICE will act as a ground support. Such an application of OWC is found in a state-of-the-art commercial e-CVT transaxle [42]. The generator starts applying positive torque to crank the ICE as soon as the EMS controller commands for all-electric to hybrid-electric mode-shift. All-electric to hybrid-electric mode-shift is highly susceptible to ICE ripple torque due to low-speed operation of ICE. In literature, numerous articles have prescribed different versions of DCCS to cope up with the adverse effect of ICE

ripple torque on transmission output, predominantly during all-electric to hybrid-electric mode-shift in parallel [184], series-parallel [199], power-split [176], compound power-split [177], and state-of-the-art parallel [188] architectures. Expression of high-frequency ICE ripple torque, a function of instantaneous ICE speed and indicated torque, is given in Eq.4.3.14 [140]. In this article, the high-frequency ICE ripple torque is replaced by low-frequency mean-value ICE resistance torque, as shown in Fig.8.13. It is also depicted in Fig.8.13.a and Fig.8.13.b that the low-fidelity ICE is modeled with the mean-value WOT torque curve and the fuel-consumption map. , where  $\tau_{fr_{ice}}$ ,  $C_{0,1,2}$ ,  $\theta_{ice}$ ,  $V_d$ ,  $\tau_{ind}(\theta_{ice})$  are friction torque of ICE, approximating factors, ICE crank angle, ICE volume, and indicated torque. During cranking, the generator takes the responsibility of providing enough positive torque to the ICE crankshaft so that the ICE speed can surpass firing speed while tussling against ICE friction torque. During cranking, traction motor compensates the rest of the responsibility in terms of sufficing torque demand at the transmission output. The expressions of coordinated torque for generator and traction motor during mode#2 to mode#1 transition is given in Eq.8.5.5.

### 8.5.6 Coordinated control for Hybrid-electric to All-electric Mode-shift

The ICE stopping/dragging phase during hybrid-electric to all-electric mode-shift is equally victimized to the effect of ICE resistance torque due to the ICE's operation below idle speed [177]. Hence, not only the cranking phase but the ICE dragging phase should also be controlled judiciously with coordinated control. Only the friction-torque component remains in the torque supplied by the ICE when the

$$\tau_{gen|coordinated}|_{2-to-1 \text{ shift}} = \frac{-\tau_{fr_{ice}} + J_{c1}\ddot{\theta}_{ice} + (\beta_1 + 1)J_{s1}\ddot{\theta}_{gen} + \frac{1}{\beta_2}(\tau_{mot|coordinated} - J_{r2}\ddot{\theta}_{mot})}{(\beta_1 + 1)} \quad (8.5.5a)$$

$$\tau_{mot|coordinated}|_{2-to-1 \text{ shift}} = \frac{\beta_2\{(\tau_{gen|coordinated} - J_{s1}\ddot{\theta}_{gen}) + \frac{(\beta_2+1)}{\beta_2}J_{r2}\ddot{\theta}_{mot} - (\tau_{out} - J_{c2}\ddot{\theta}_{out})\}}{(\beta_2 + 1)} \quad (8.5.5b)$$

$$\tau_{gen|coordinated}|_{4-to-3 \text{ shift}} = J_{s1}\ddot{\theta}_{gen} - \frac{1}{1 + \beta_1}(\tau_{fr_{ice}} - J_{c1}\ddot{\theta}_{ice}) + \frac{1}{\beta_2}(\tau_{mot|coordinated} - J_{r2}\ddot{\theta}_{mot}) \quad (8.5.6a)$$

$$\tau_{mot|coordinated}|_{4-to-3 \text{ shift}} = \beta_2 \left\{ \frac{\frac{\beta_2+1}{\beta_2}J_{r2}\ddot{\theta}_{mot} - \frac{\beta_1}{1+\beta_1}(\tau_{fr_{ice}} - J_{c1}\ddot{\theta}_{ice}) - (\tau_{out} - J_{c2}\ddot{\theta}_{out})}{\beta_2 + 1} \right\} \quad (8.5.6b)$$

EMS controller commands for fuel cut-off of the ICE. Although the EMS commands the powertrain to be operated in a all-electric mode, the powertrain cannot be operated in all-electric mode until  $\omega_{ice}$  reduces to zero. The expressions for coordinated torque of generator and traction motor remain the same as Eq.8.5.5 and Eq.8.5.6. Reduction of  $\omega_{ice}$  is expedited by applying negative torque through the generator and the powertrain can be technically shifted to all-electric mode as soon as  $\omega_{ice}$  reduces to zero.

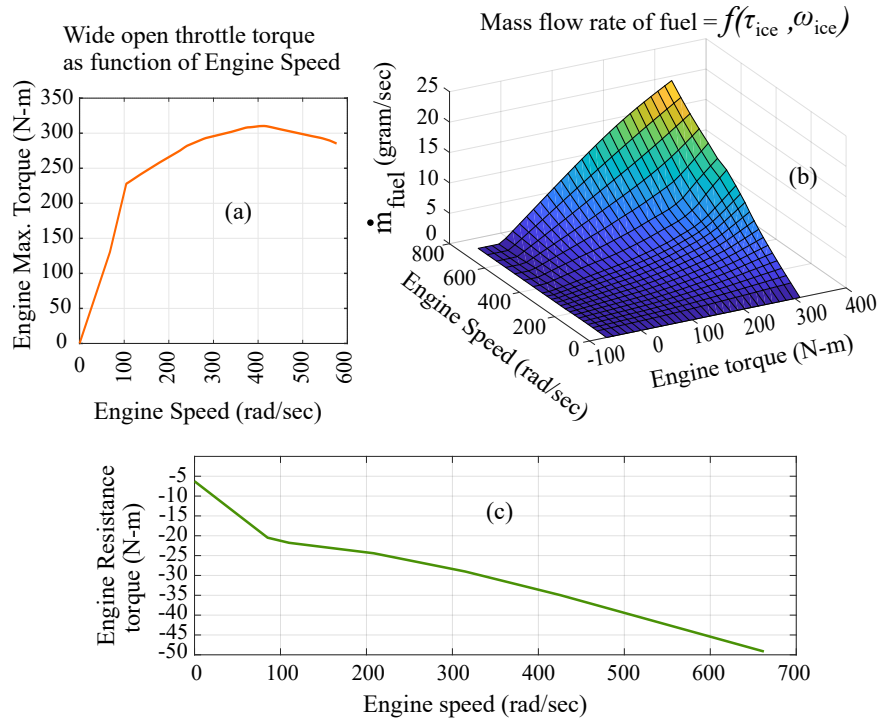


Figure 8.13: ICE characteristics curves required for modeling the ICE. (a) Mean-value wide open throttle torque as a function of ICE speed. (b) Mass-flow rate of fuel consumption as a function of ICE speed and torque. (c) Mean-value ICE resistance torque as a function of ICE speed.

## 8.6 Instantaneous Optimization-based Mode Scheduling Control for Multi-mode ePT

ECMS is a lucrative option as a control strategy for an EMS dealing with a multi-mode electrified powertrain [15]. The ability of real-time implementation accentuates ECMS as a strong contender for optimal utilitarian control strategy securing charge sustainability along with minimizing fuel consumption.

The selection of powertrain operating mode (among the four feasible modes) through a two-stage optimization process, as given in Eq.8.6.1, is the conspicuous identity of the multi-mode ECMS. Such a selection process is supervised by the comparison of

equivalent operating cost among all the four modes at every time-step. The equivalent cost structure used by multi-mode ECMS is comprised of two components, i.e., the actual chemical energy emanated from combustible fuel and equivalent chemical energy representing the electrical energy from HVB.

$$[u_{mode}^{opt}(t)]_{\#i} = \arg \min_{u^{(k)}_{mode\#i}} [(J)_{equi.}], \quad t = 1, \dots, T_f \quad (8.6.1a)$$

$$[u_{mode}^{opt}(t)] = \arg \min_{u_{mode\#i}(t)} [(J)_{equi.}], \quad t = 1, \dots, T_f \quad (8.6.1b)$$

$$(J)_{equi.} = \dot{m}_{fuel} \times (LHV)_{fuel} + \lambda \times Pow_{batt}. \quad (8.6.2)$$

The equivalent cost structure is given in Eq.8.6.2, where  $\dot{m}_{fuel}$  is the mass flow rate of combustible in (gm/sec),  $LHV$  is the lower heating value of the fuel in (kJ/gm),  $\lambda$  is the EF between stored electrical energy in HVB and stored chemical energy in the fuel, and  $u_{control}$  represents the set of control actions corresponding to a particular mode.

The control actions corresponding to each of the four modes are delineated in Tab.8.2. A utilitarian aspect in the form of drivability improvement can be augmented in the list of objectives to make the multi-mode ECMS apposite for practical implementation. Multi-mode ECMS has a propensity to change the powertrain mode too often, and it can quickly deteriorate the drivability, i.e., driver's perception about driving comfort. Hence, it is highly advised to add a mode-shift penalty in the cost structure to prevent multi-mode ECMS from yielding excessive mode-shifts. Inspired by the work presented in [142], the original equivalent cost structure, given in Eq.8.6.2, has

been modified as follows:

$$(J)_{equi}^{modif.} = \dot{m}_{fuel} * (LHV) + \lambda * Pow_{batt.} + \nu * P_{shift} \quad (8.6.3)$$

, where

$$\begin{aligned} P_{shift_{mode}} &= \mu_1 * \frac{1}{2} J_{ice} [(\omega_{ice}^{t+1})^2 - (\omega_{ice}^t)^2] \\ &+ \mu_2 * \frac{1}{2} J_{gen} [(\omega_{gen}^{t+1})^2 - (\omega_{gen}^t)^2] \\ &+ \mu_2 * \frac{1}{2} J_{mot} [(\omega_{mot}^{t+1})^2 - (\omega_{mot}^t)^2] \end{aligned}$$

$\nu$  is a tuning factor, and a value of 0.03 is selected for this article. Different values have been assigned to tuning factors  $\mu_1$ , and  $\mu_2$  corresponding to different types of mode-shift and the assigned values are tabulated in Tab.8.4. Traditional ECMS generates the mode-shift schedule based only on the steady-state mode operations' powertrain dynamics and not on the powertrain dynamics during transient mode-shift events. In this article, the authors have proffered an ECMS, assisted by the coordinated control, for a real-time EMS that yields optimal mode schedule while considering transients and feasibility of the mode-shift events. It has been reiterated on copious occasions in literature that EF plays an essential role in the conservation of charge sustainability and optimality of the yielded control from ECMS. The value of EF ( $\lambda$ ) strictly governs the selection of mode at any given time-step. A constant optimal value of EF for the entire drive cycle can be obtained through trial-and-error method or iteratively, given that the drive cycle is known in advance. Since this article elucidates the effect of coordinated control on optimal mode selection of multi-mode ECMS, and since it does not focus on proffering a disruptive derivative of ECMS yielding optimal performance

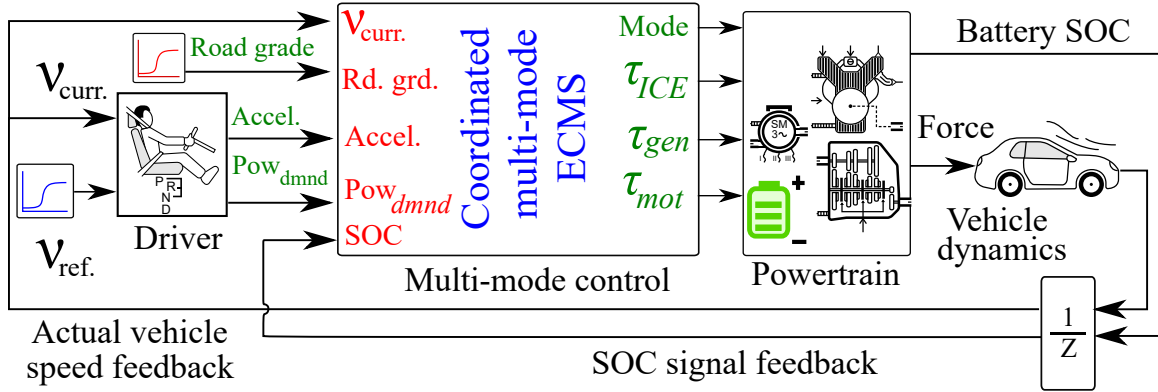


Figure 8.14: Implementation of coordinated control in a multi-mode ePT where mode scheduling is governed by ECMS for unknown drive cycles, the authors are privileged to access the prior knowledge of the drive cycle.

## 8.7 Simulation Results and Discussion

In order to verify the effectiveness of the proposed mode scheduling control, assisted by coordinated control, a vehicle simulation model is developed in Simulink (Mathworks®) platform, as depicted in Fig.8.14. The specifications of the vehicle components and powertrain components used in the OSP are furnished in Tab.4.1.

Validation of the vehicle plant model is an integral part of the optimal mode-scheduling control development. Since the powertrain topology is generated from ATST and no hardware prototype is available for physical testing, experimental validation is out of scope, at least for the current article. Now, the comprehensive objective of this simulation framework should be reiterated here. The online simulation framework's objective is to certify whether the new optimal topology candidates should be delegated to the next level of attestation (i.e. HIL) before final approval for hardware prototyping. Now, since the framework presented here itself chronologically appears



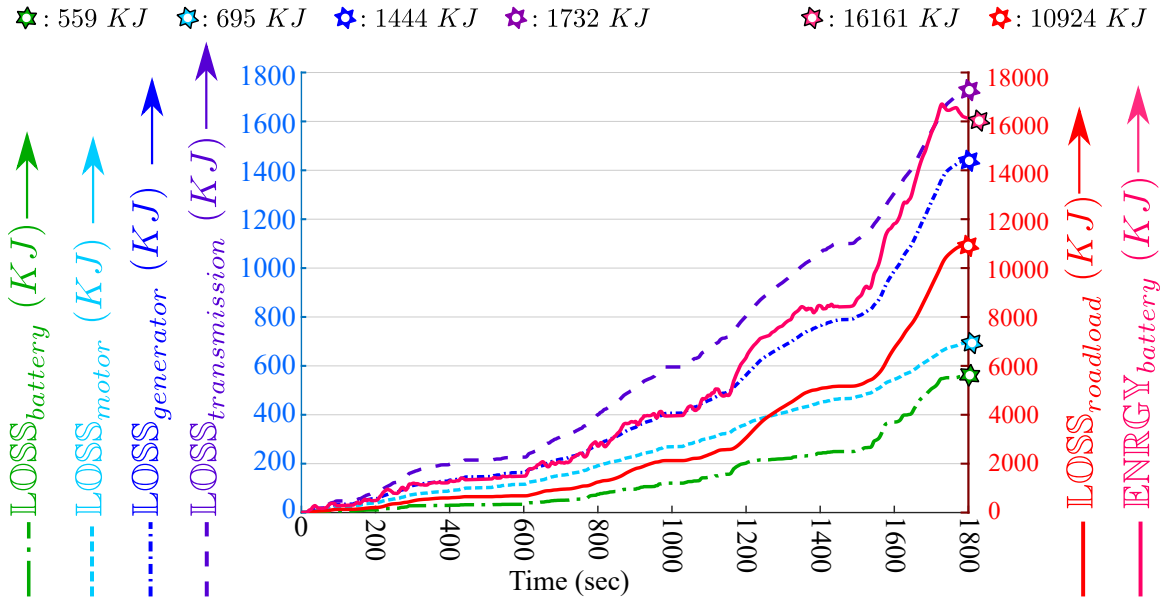


Figure 8.15: Trajectories of major components' energy and major losses in 1<sup>st</sup> All-electric mode (Mode#3) under WLTC drive cycle. The total energy of HVB and the total losses incurred by the HVB, generator, motor, transmission, and road load at the end of the drive cycle are furnished here to support the energy balance presented in Tab.8.5.

long before hardware prototyping, experimental validation will not be a justified demand. Hence, there is a lucid necessity of a theoretical validation of the vehicle plant model conflated with the EMS controller. Energy balance, which is defined as the summation of all lost energies and supplied energies, has been a well-accredited method for validating vehicle simulation models. Tab.8.5 furnishes energy balance for two all-electric modes and one hybrid-electric mode under a standard WLTC. In ideal modeling condition the last row of Tab.8.5 should be zero for all three modes. However, errors yielded in all the modes are well under permissible limit which is 0.1% of total energy spent throughout the drive cycle. Fig.8.15 depicts how the cumulative energy supplied by the HVB (high-voltage battery) increases throughout the drive cycle in the 1<sup>st</sup> all-electric mode. Fig.8.15 also depicts how the cumulative energy

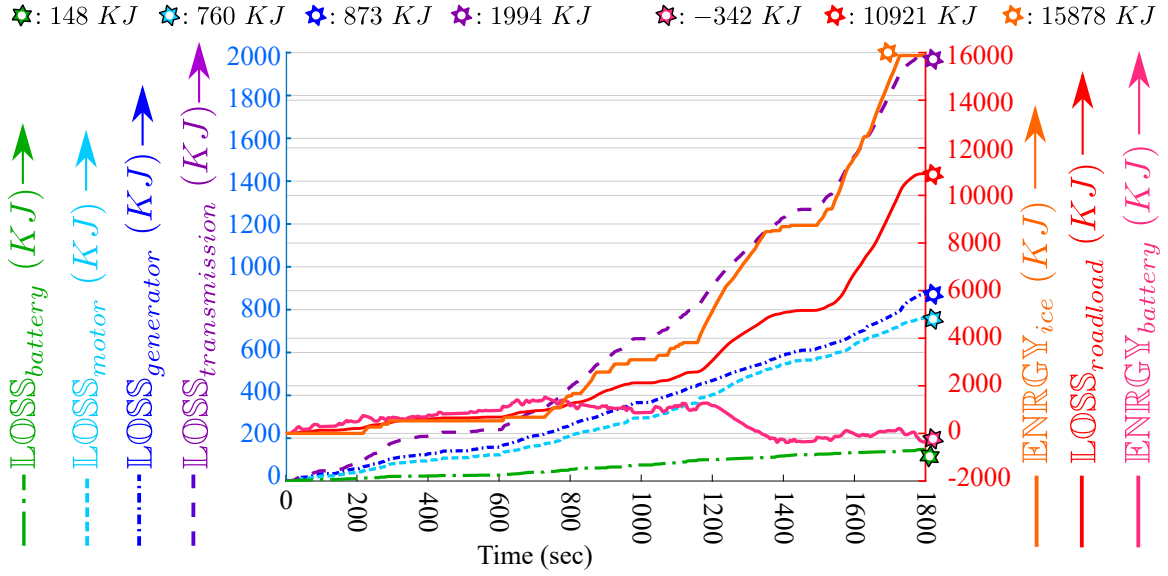


Figure 8.16: Trajectories of major components' energy and major losses in 1st Hybrid-electric mode (Mode#1) under WLTC drive cycle. The total energy supplied by the ICE and HVB and the total losses incurred by the HVB, generator, motor, transmission, and road load at the end of the drive cycle are furnished here to support the energy balance presented in Tab.8.5.

lost in different loss-incurring components increases throughout the drive cycle in the 1<sup>st</sup> all-electric mode. *Road load loss, transmission loss, motor loss, generator loss, and HVB loss* are the significant consumers of energy solely provided by the HVB output in the 1<sup>st</sup> all-electric mode, as shown in Eq.8.7.1.

$$\begin{aligned}
 \left( \text{ENERGY}_{\text{battery}} \right)_{\text{all-electric mode}} &= \text{LOSS}_{\text{roadload}} \\
 &+ \text{LOSS}_{\text{transmission}} + \text{LOSS}_{\text{generator}} \\
 &+ \text{LOSS}_{\text{motor}} + \text{LOSS}_{\text{battery(HVB)}}
 \end{aligned} \tag{8.7.1}$$

Fig.8.15 indicates that the total useful energy provided by the HVB at the end of the drive cycle is 16161 KJ, and the total energies lost through *road load loss, transmission loss, motor loss, generator loss, and HVB loss* are 10924 KJ, 1732 KJ, 695 KJ,

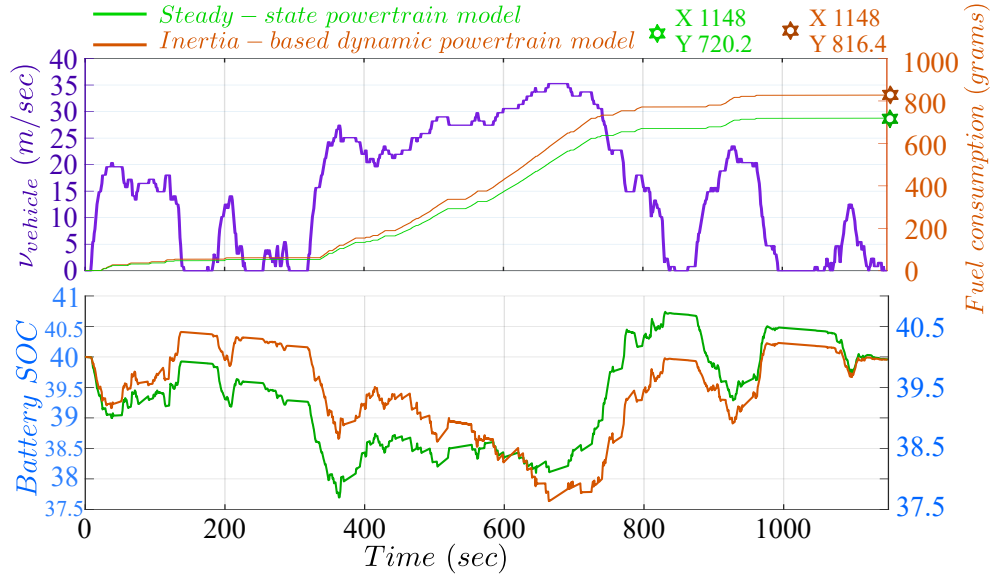


Figure 8.17: Comparison of fuel consumption and battery SOC trajectories between inertia-based and steady-state ePT models simulated with same EMS strategy.

1444  $KJ$ , and 559  $KJ$ , respectively.

Similarly, Fig.8.16 supports the energy balance of a hybrid-electric mode, as given in the 5<sup>th</sup> column of Tab.8.5 by indicating total useful energies supplied by ICE and HVB throughout the drive cycle, and by indicating total energies lost through *road load loss*, *transmission loss*, *motor loss*, *generator loss*, and *HVB loss*. The energy balance in a hybrid-electric mode can be formulated as Eq.8.7.2.

$$\begin{aligned}
 & \left( \text{ENRGY}_{\text{battery}} + \text{ENRGY}_{\text{ice}} \right)_{\text{hybrid-electric mode}} = \\
 & \text{LOSS}_{\text{roadload}} + \text{LOSS}_{\text{transmission}} \\
 & + \text{LOSS}_{\text{generator}} + \text{LOSS}_{\text{motor}} + \text{LOSS}_{\text{battery(HVB)}}
 \end{aligned} \tag{8.7.2}$$

After the vehicle simulation platform gets validated, it is time to show the result of the optimal mode schedule obtained through multi-mode ECMS, which is assisted by coordinated control. A real-world driving mission, as shown in Fig.8.17, has been

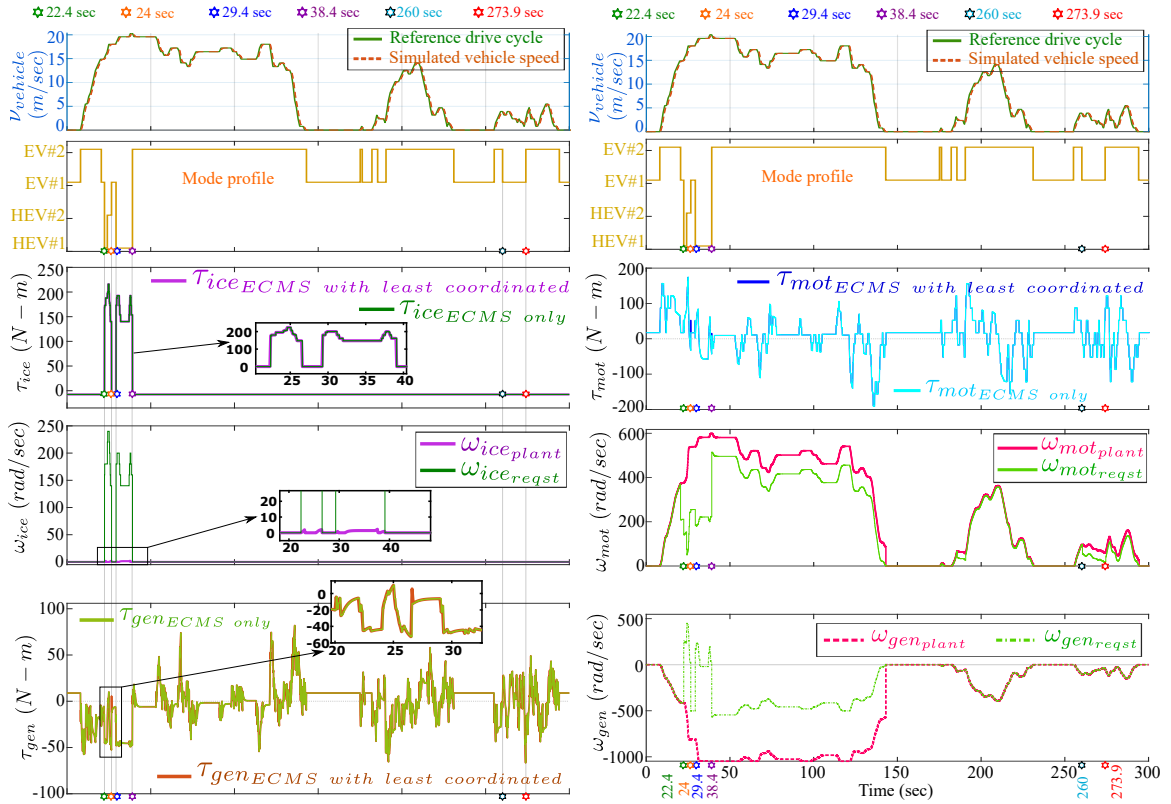


Figure 8.18: Timeseries of vehicle speed,  $\omega_{ice}$ ,  $\omega_{mot}$ ,  $\omega_{gen}$  as a result of ECMS without any coordinated control.

chosen for demonstrating the effectiveness of the proposed control strategy for multi-mode ePT.

Before advancing with the demonstration, it is crucial to reiterate the significance of an inertia-based dynamic model of a multi-mode ePT over a steady-state model. Unlike the inertia-based ePT model, the steady-state model does not account for the transient dynamics of the powertrain components, and therefore steady-state model incurs lesser loss than the inertia-based model [201]. Consequently, the steady-state model will consume lesser fuel than the inertia-based model when both ePT models are simulated for the same drive cycle under the same EMS controller, as shown in Fig.8.17. The fuel consumption is measured under the charge sustainable

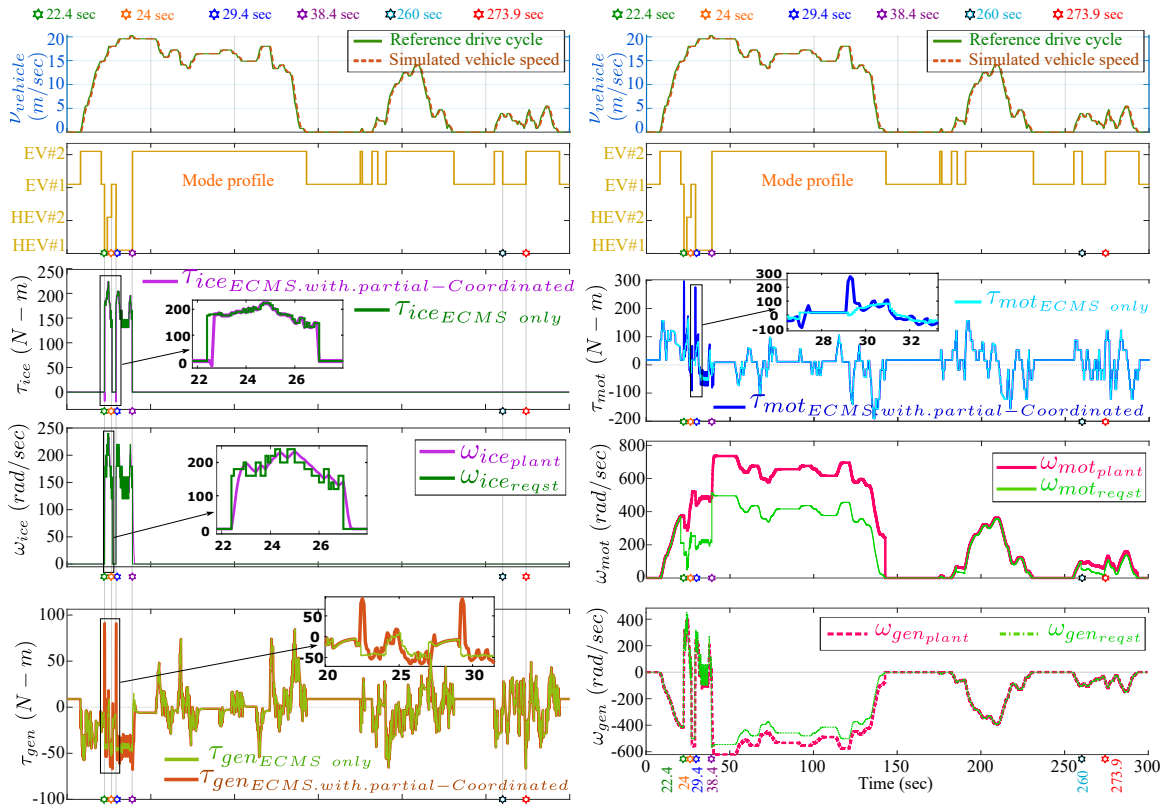


Figure 8.19: Timeseries of vehicle speed,  $\omega_{ice}$ ,  $\omega_{mot}$ ,  $\omega_{gen}$  as a result of ECMS assisted with partial coordinated control, i.e., during EV-to-HEV and HEV-to-EV mode shifts.

condition in both models for a fair comparison. The inertia-based dynamic model will emulate a real HEV's performances, such as fuel-economy, charge sustainability, and drivability, more pragmatically than a steady-state model. As far as the screening and performance evaluation of optimal ePT configurations, which will be prototyped in the future, is concerned, why wouldn't someone choose a pragmatic inertia-based ePT model over the steady-state model?

Although the choice of a pragmatic inertia-based model seems obvious, its implementation requires a significant improvement of the vehicle plant model and the assistance of a separate coordinated control for the EMS controller. An EMS control that is sufficient for a steady-state ePT model will not be sufficient alone for an inertia-based

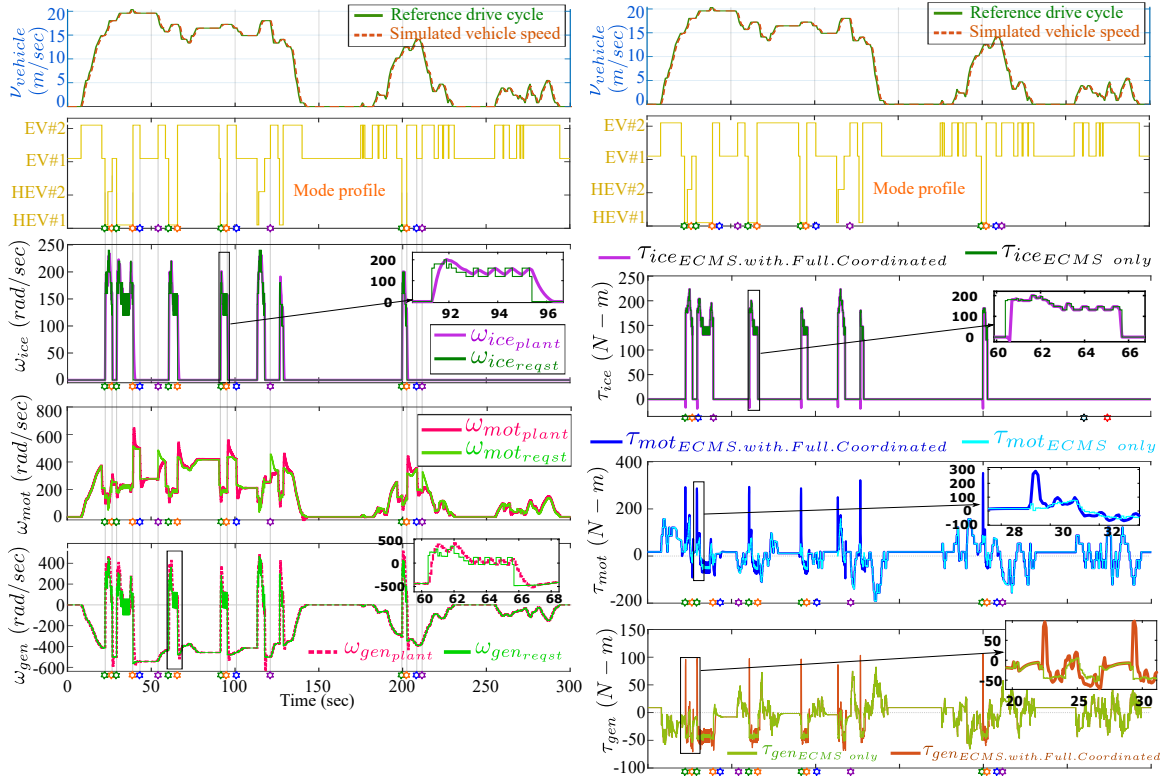


Figure 8.20: Timeseries of vehicle speed,  $\omega_{ice}$ ,  $\omega_{mot}$ ,  $\omega_{gen}$  as a result of ECMS assisted with full coordinated control, i.e., during EV-to-HEV, HEV-to-EV, HEV-to-HEV, and EV-to-EV mode shifts.

ePT model. The coordinated control crucially modifies the EMS control during mode-shift events so that the inertia-based dynamic ePT model can emulate a real physical ePT. The articulation of a coordinated control for an ePT highly depends on the ePT's configuration. The articulation stages are more straightforward if the ePT has only one all-electric mode and another hybrid-electric mode of operation. The articulation stages of the overall coordinated control become convoluted as the ePT configuration offers multiple all-electric and hybrid-electric modes. The next paragraphs will delineate the results obtained throughout different stages of coordinated control implementation for a multi-mode ePT, and will corroborate coordinate control's effectiveness in assisting the ECMS controller to choose the correct mode of

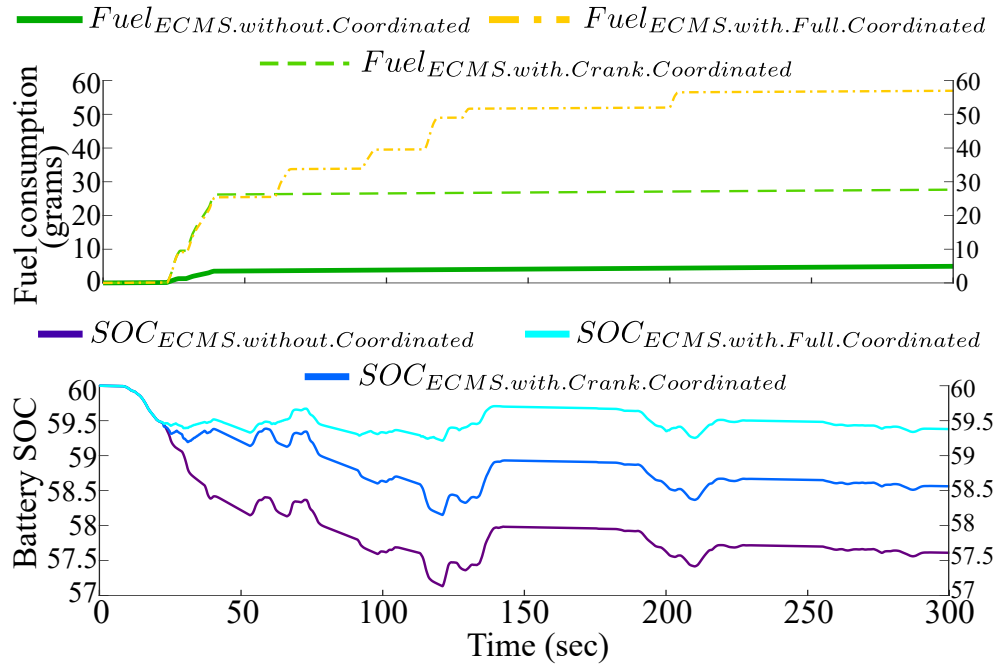


Figure 8.21: Comparison of SOC and fuel consumption trajectories among three aforementioned cases, i.e., ECMS without coordinated control, ECMS with partial coordinated control, and ECMS with full coordinated control.

operation.

For a better demonstration, the first 300 seconds of the real-world drive cycle instead of the entire drive cycle, as shown in Fig.8.17, is simulated separately with three coordinate control implementation stages. The simulation platform is comprised of two main parts, i.e., the supervisory controller (EMS), equipped with different stages of coordinated controlled ECMS, and the vehicle plant model. The inertia-based powertrain dynamics, as expressed in Eq.8.4.2, Eq.8.4.4, Eq.8.4.6, and Eq.8.4.8 is the heart of the vehicle plant model, and inputs to the powertrain dynamics model are the controlled torque signals ( $\tau_{ice}$ ,  $\tau_{mot}$ ,  $\tau_{gen}$ ) and mode signal from the supervisory controller, and outputs of the powertrain dynamics model are the angular acceleration of ICE, traction motor, generator, and wheel. It is noteworthy to mention that

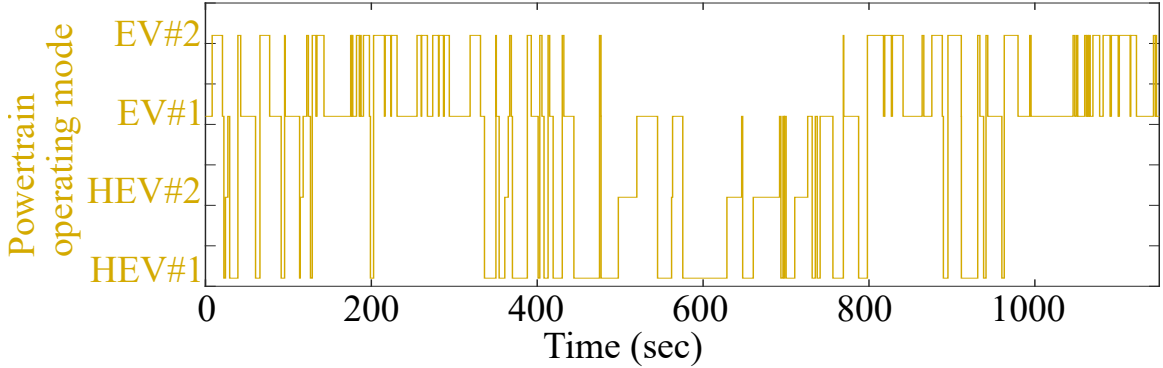


Figure 8.22: Time-series of optimal powertrain operating mode yielded by ECMS assisted by full coordinated control for an entire real-world drive cycle.

the EMS does not cater to the powertrain dynamics model explicitly with angular speed commands of the aforementioned prime movers but internally computes the angular speed values. The success of coordinated control depends on whether the coordinated torque signals can yield the same angular speed of the prime movers that was internally computed and requested by the EMS. Hence, a comparison between requested angular speed and simulated angular speed will corroborate the coordinated control’s efficacy throughout the drive cycle.

At the first stage, the vehicle plant model has been actuated with control inputs spewed directly from ECMS, i.e., without assistance from coordinated control, to observe the plant model’s responses in terms of  $\nu_{vehicle}$ ,  $\omega_{ice}$ ,  $\omega_{mot}$ , and  $\omega_{gen}$ . The control inputs and plant responses are depicted in Fig.8.18. The 2<sup>nd</sup> row of Fig.8.18 shows that ECMS commands for a mode-shift from 1<sup>st</sup> all-electric (1<sup>st</sup> EV) to 1<sup>st</sup> hybrid-electric (1<sup>st</sup> HEV) at 22.4<sup>th</sup> second, and ECMS commands for an 2<sup>nd</sup> HEV to 1<sup>st</sup> EV mode-shift at 24<sup>th</sup> second. Again there is a command of an 1<sup>st</sup> EV to 1<sup>st</sup> HEV mode-shift at 29.4<sup>th</sup> second and 1<sup>st</sup> HEV to 2<sup>nd</sup> EV mode-shift at 38.4<sup>th</sup> second. The 3<sup>rd</sup> row of Fig.8.18 shows the exact resemblance between  $\tau_{ice_{ECMS-without-coordinated}}$  and



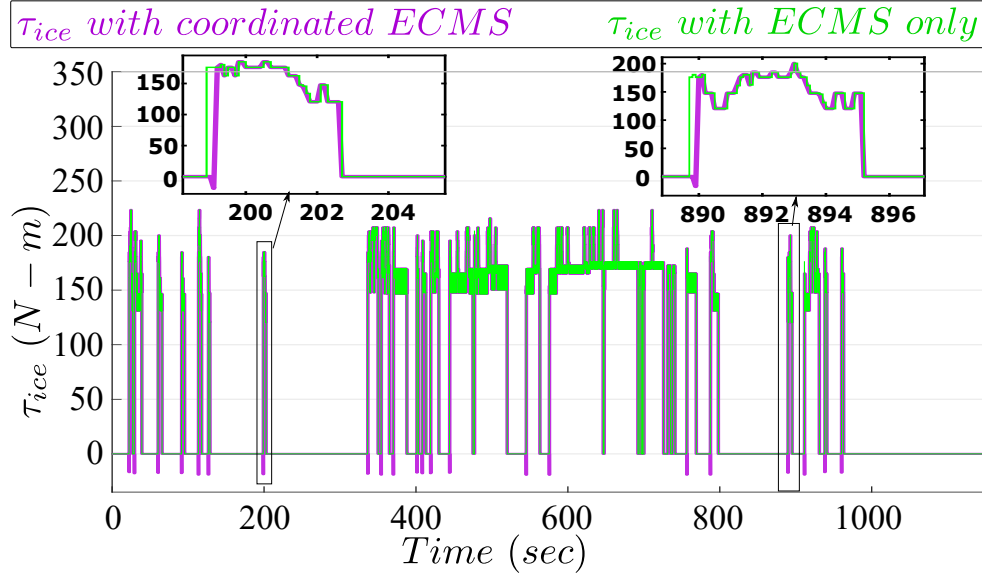


Figure 8.23: Comparison between time-series of  $\tau_{ice_{optim.}}$  yielded by “full coordinated control assisted ECMS” and ECMS only for an entire real-world drive cycle.

$\tau_{ice_{ECMS-only}}$ . The 5<sup>th</sup> and 7<sup>th</sup> rows of Fig.8.18 show same trend in terms of exact resemblance of  $\tau_{mot_{ECMS-without-coordinated}}$  with  $\tau_{mot_{ECMS-only}}$  and  $\tau_{gen_{ECMS-without-coordinated}}$  with  $\tau_{gen_{ECMS-only}}$ , respectively. The inertia-based ePT model’s responses to the torques mentioned above are given in the 1<sup>st</sup>, 4<sup>th</sup>, 6<sup>th</sup>, and 8<sup>th</sup> rows of Fig.8.18. It is lucid that the  $\nu_{vehicle}$  perfectly correlates with the reference drive cycle despite the absence of coordinated control. However,  $\omega_{ice_{plant}}$  barely reaches 5 rad/sec and could not follow the  $\omega_{ice_{rqst}}$  due to the absence of any coordinated control during 1<sup>st</sup>EV to 1<sup>st</sup> HEV mode-shift. Similarly,  $\omega_{mot_{plant}}$  and  $\omega_{gen_{plant}}$  could not follow  $\omega_{mot_{rqst}}$  and  $\omega_{gen_{rqst}}$ , respectively during the all-electric to hybrid-electric mode-shifts and vice-versa. The mismatch between requested speed and plant speed for both traction motor and generator persisted even after the termination of 1<sup>st</sup> HEV mode at 38.4<sup>th</sup> second and throughout the 2<sup>nd</sup> EV mode’s operation until the vehicle’s speed is reduced to zero around 150<sup>th</sup> second. The reason of such mismatch is the absence of coordinated control during HEV-to-EV mode-shift. There is no mismatch between  $\omega_{gen_{plant}}$  and

$\omega_{gen_{rqst}}$  after 150<sup>th</sup> second. However, there are mismatches between  $\omega_{mot_{plant}}$  and  $\omega_{mot_{rqst}}$  despite all-electric operation. At 260<sup>th</sup> second, there is a command for 2<sup>nd</sup> EV to 1<sup>st</sup> EV mode-shift while the vehicle speed is 4 *m/sec*. Since there is no coordinated control during EV-to-EV mode-shifts,  $\omega_{mot_{plant}}$  could not correlate with the  $\omega_{mot_{rqst}}$  if vehicle speed is greater than zero during EV-to-EV shifts.

At the second stage, ECMS is assisted by the coordinated control during EV-to-HEV and HEV-to-EV mode-shifts. The controlled torques from coordinated-ECMS and the responses from the vehicle plant model are shown in Fig.8.19. 3<sup>rd</sup>, 5<sup>th</sup>, 7<sup>th</sup> rows of the Fig.8.19 depict a lucid change in the trajectories of  $\tau_{ice}$ ,  $\tau_{mot}$ , and  $\tau_{gen}$ . The coordinated control enables the generator to churn some extra torque, which helps the ICE surpass its idle speed. Hence, the coordinated control should adjust the motor torque during the ICE cranking phase to ensure appropriate torque supply at the transmission output. The 4<sup>th</sup> row of Fig.8.19 corroborates the rudimentary coordinated control's effectiveness through a decent correlation between  $\omega_{ice_{plant}}$  and  $\omega_{ice_{rqst}}$ . Compared to Fig.8.18, Fig.8.19 shows partial, but not complete, improvement in correlation between  $\omega_{gen_{plant}}$  and  $\omega_{gen_{rqst}}$  due to rudimentary coordinated control. Slight mismatch yielded during 1<sup>st</sup> HEV to 2<sup>nd</sup> EV mode-shift persists throughout the 2<sup>nd</sup> EV mode. Absence of coordinated control during HEV-to-HEV mode-shift can be held accountable for the abovementioned mismatch and the mismatch between  $\omega_{mot_{plant}}$  and  $\omega_{mot_{rqst}}$ .

Finally, the performance of an ECMS with full-fledged coordinated control is presented in Fig.8.20. The final version of coordinated control is well equipped to tackle all types, i.e., EV-to-HEV, HEV-to-EV, HEV-to-HEV, and EV-to-EV, mode-shift dynamics, and to yield praiseworthy correlation between the requested speed and plant

speed for all the prime-movers. It is noteworthy to mention that the mode profile has changed significantly after employing the final version of coordinated control, assisting the optimal ECMS to decide optimal mode-shift events. ECMS's mode-selection algorithm has a critical reliance on the feedback signals of the prime-movers. Hence, the accuracy of the feedback signals plays a crucial role in the mode-selection algorithm. The final version of coordinated control's effectiveness can be further unveiled by comparing fuel consumption and HVB's SOC trajectory among all three versions of coordinated control, as shown in Fig.8.21. Fig.8.21 corroborates the impact of coordinated control's articulation on the accuracy of a multi-mode ePT's performance evaluation. It is also noteworthy to mention that vehicle's speed is very well correlated regardless of the version of coordinated control. Hence, a coordinated control's competency should be judged based on its ability to correlate all prime mover's speed instead of just vehicle's speed. An incompetent coordinated control will eventually evaluate the performance of a multi-mode ePT inaccurately, which will lead to misidentification of optimal configurations with multi-mode ePT. The result section culminates with demonstrating the performance of the ECMS assisted by coordinated control for an entire drive cycle. A constant value of 2.695 is used for the EF for satisfying charge sustenance. The optimal mode profile, as shown in Fig.8.22, has been found through an iterative process that is dedicated to finding a constant EF sustaining the battery SOC. The value of EF predominantly governs the mode selection inside the ECMS-based supervisory controller, whereas coordinated control modifies the control torque signals during mode-shift events. Fig.8.23, Fig.8.24, and Fig.8.26 depict the effect of coordinated control on conventional ECMS during mode-shift events. Catered with these coordinated torque signals, the powertrain dynamics model emits the simulated

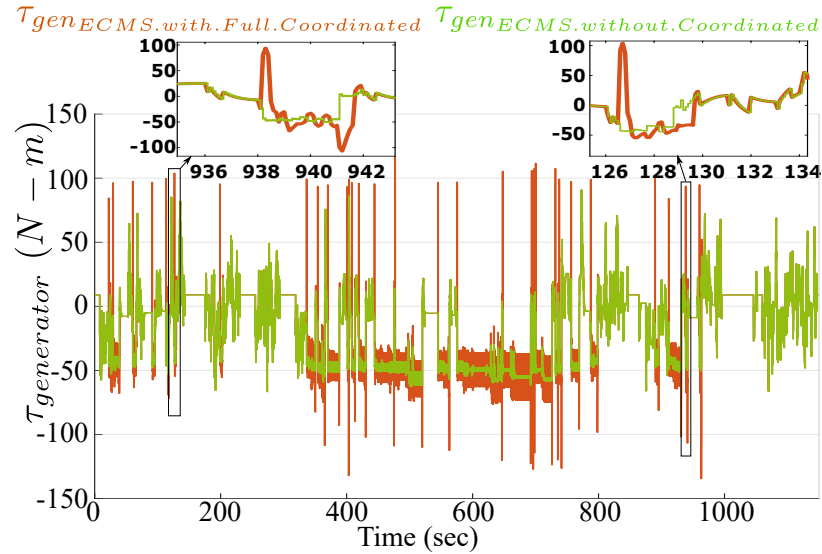


Figure 8.24: Comparison between time-series of  $\tau_{gen_{optim}}$  yielded by “full coordinated control assisted ECMS” and ECMS only for an entire real-world drive cycle.

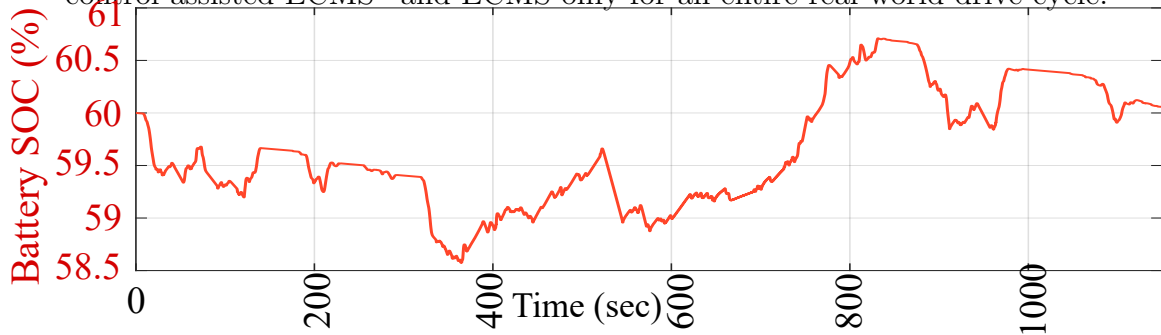


Figure 8.25: Battery SOC trajectory yielded with ECMS assisted by full coordinated control for an entire real-world drive cycle.

$\omega_{ice}$ ,  $\omega_{gen}$ ,  $\omega_{mot}$ , and  $\omega_{wheel}$ . Fig.8.27, Fig.8.28, and Fig.8.29 depict that the coordinated control assists the ICE, generator, and traction motor respectively in achieving the requested speeds throughout the drive cycle, especially during the mode shifts. Although the effectiveness of coordinated control has already been corroborated with the comparative study among Fig.8.18, Fig.8.19, and Fig.8.20, Figs.8.27,8.28, and 8.29 are furnished to showcased performance of coordinated ECMS throughout a complete drive cycle.

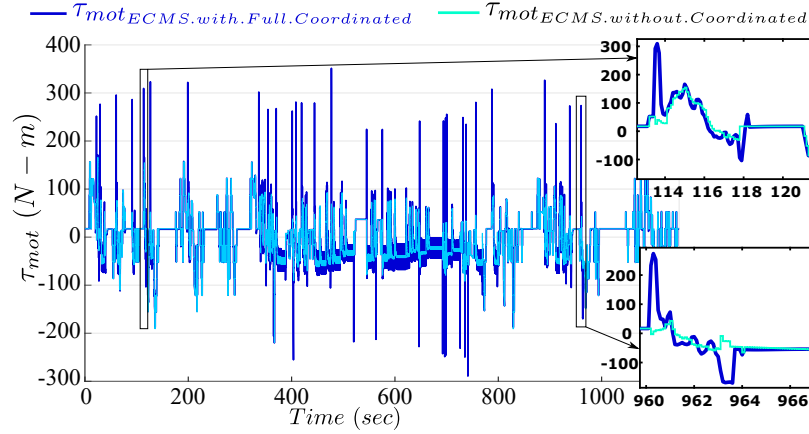


Figure 8.26: Comparison between time-series of  $\tau_{mot_{optim}}$  yielded by “full coordinated control assisted ECMS” and ECMS only for an entire real-world drive cycle.

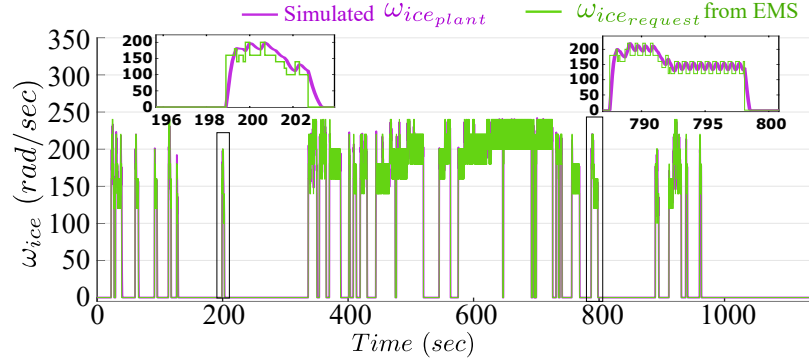


Figure 8.27: Correlation between ICE speeds yielded from vehicle plant catered with “full coordinated control assisted ECMS” and the speed requested by EMS.

## 8.8 Summary

This chapter proffers four essential contributions in the domain of multi-mode electrified powertrain. It streamlines the derivation of inertia-based dynamic equations for all the feasible modes facilitated by the powertrain topology. The dynamics associated with each of the feasible mode-shift events are modeled here. It corroborates the incumbency of a coordinated control for obtaining an optimal mode schedule. Finally,

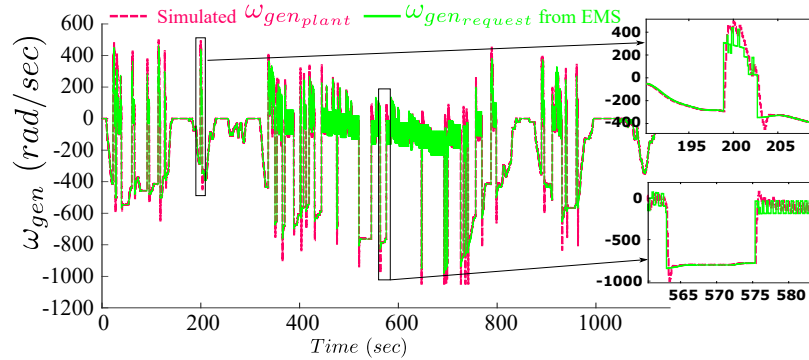


Figure 8.28: Correlation between generator speeds yielded from vehicle plant catered with “full coordinated control assisted ECMS” and the speed requested by EMS.

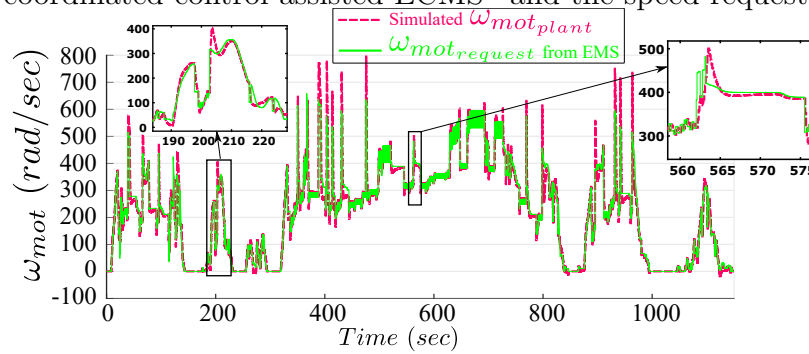


Figure 8.29: Correlation between traction motor speeds yielded from vehicle plant catered with “full coordinated control assisted ECMS” and the speed requested by EMS.

the article proposes the effectiveness of coordinated control in the powertrain configuration’s correct performance evaluation. At the end, an equivalent consumption minimization strategy-based mode-scheduling strategy, assisted by the coordinated control is applied in the energy management system for obtaining optimal mode profile for a real-world drive cycle for corroborating the efficacy of the coordinated control during all feasible mode-shift events.

Considering the fact that the article focuses on system level performance evaluation, the coordinated control is not articulated for eradicating high-frequency ICE ripple torque, instead, it is designed to tackle mean-valued ICE resistance torque during

ICE cranking and fuel cut-off. The powertrain dynamics has been numerically validated through an *Energy balance* method instead of an experimental validation due to any prototype or commercial vehicle's unavailability with the concerned multi-mode electrified powertrain. However, experimental validation is planned for future work. This article will distinctively stand out from the literature on coordinated control of multi-mode electrified powertrain based on its contribution to system level assessment rather than component level.

The entire online simulation framework presented in this article can be institutionalized as the first step of the performance validation for a new topology of multi-mode electrified powertrain. The article paves the way for a new research direction where coordinated control can be applied simultaneously to the system level to affect the mode scheduling strategy and the component level to neutralize the high-frequency perturbation generated from the ICE ripple torque.

Table 8.1: Specification of the vehicle and cardinal components

Component	Parameter	Value
Vehicle	Mass (m)	2510 kg
	Rolling resistance co-efficient	$f_1 = 0.006, f_2 = 0.0001$
	Vehicle inertia ( $J_{out}$ )	298.08 kg-m <sup>2</sup>
	Wheel radius ( $r_{wh}$ )	0.3548 m
	Vehicle height	1.75 m
	Vehicle width	1.48 m
	Engine	Inertia ( $J_{ice}$ )
Compression ratio		
Max. speed ( $\omega_{ice_{max}}$ )		5500 RPM
Max. torque ( $\tau_{ice_{max}}$ )		310 Nm at 4100 rpm
Volume displacement		4 Cyl., 2.0152 liters
Max. power		175 kWatt at 5000 rpm
Traction motor	Inertia ( $J_{mot}$ )	0.0028 kg-m <sup>2</sup> ,
	Max. speed, Max. Volt.	12500 RPM, 414 V
	Max. torque ( $\tau_{mot_{max}}$ )	410 Nm at [0:2240]
	Max. power	96 kWatt at 2240 rpm
Generator	Inertia ( $J_{gen}$ )	0.0024 kg-m <sup>2</sup>
	Max. speed, Max. Volt.	12500 RPM, 350 V
	Max. torque ( $\tau_{gen_{max}}$ )	162.5 at [0:4167] rpm
	Max. power	70 kWatt at 4167 rpm
Battery	Max. capacity	43 Ah
	Max. discharge power	60 kWatt
	Max. recharge power	40 kWatt
	No. of cells in series	100
	Max. voltage	414 V

Table 8.2: Description of modes facilitated by the new electrified powertrain topology

Mode	Clutch#1	Clutch#2	Clutch#3	DoF ( $u_{control}$ )
$Mode_1$	Engaged	Disengaged	Disengaged	2 ( $\tau_{ice}, \omega_{gen}$ )
$Mode_2$	Engaged	Disengaged	Engaged	1 ( $\tau_{mot}$ )
$Mode_3$	Disengaged	Engaged	Disengaged	2 ( $\tau_{ice}, \omega_{gen}$ )
$Mode_4$	Disengaged	Engaged	Engaged	1 ( $\tau_{mot}$ )



Table 8.3: Required clutch engagement and disengagement for different mode-shifts  
(infeasible mode-shifts are shown in red color)

Current Past mode mode	1 <sup>st</sup> HEV	1 <sup>st</sup> EV	2 <sup>nd</sup> HEV	2 <sup>nd</sup> EV
1 <sup>st</sup> HEV	No shift	Immed. fuel-cut, CL3 engage after $\omega_{ice} = 0$	(CL1 diseng., + CL2 engag.)	Immed. fuel-cut; (CL2 engag. + CL1 diseng.); CL3 enga. after $\omega_{ice} = 0$
1 <sup>st</sup> EV	CL3 disengag.	No shift	(CL3 diseng. + CL2 enga. + CL1 diseng.)	(CL1 diseng. + CL2 enga.)
2 <sup>nd</sup> HEV	(CL2 diseng. + CL1 enga.)	Immed. fuel-cut; (CL2 diseng. + CL1 enga.); CL3 enga. after $\omega_{ice} = 0$	No shift	Immed. fuel-cut; CL3 enga. after $\omega_{ice} = 0$
2 <sup>nd</sup> EV	(CL3 diseng. + CL1 enga. + CL2 diseng.)	(CL2 diseng. + CL1 enga.)	Only CL3 diseng.)	No shift

Table 8.4: Tuning factor for mode-shift penalty term

Past mode	Current mode	
	Fully electric	Hybrid electric
Fully electric	$\mu_1 = 0, \mu_2 = 1$	$\mu_1 = 0.1, \mu_2 = 0.5$
Hybrid electric	$\mu_1 = 0.1, \mu_2 = 0.5$	$\mu_1 = 1, \mu_2 = 1$

Table 8.5: Energy balance as measure of validating plant model

Energy balance	Powertrain Mode	1st All -electric	2nd All -electric	Hybrid -electric
Cumulative mechanical and electrical losses	Generator loss (KJ)	1444	1256	873
	Motor loss (KJ)	695	891	760
	Auxiliary loss (KJ)	864	864	864
	Fric. brake loss (KJ)	0	0	0
	Road load loss (KJ)	10924	10880	10921
	Transmission loss (KJ)	1732	1744	1994
ICE statistics	Battery loss (KJ)	559	586	148
	Total fuel consumption (grams)	0	0	1019
	Total mechanical energy supplied by ICE (KJ)	0	0	15878
Battery statistics	SOC at the start%	60	60	60
	SOC at the end%	34.87	35.56	60.52
	Extra energy stored in battery (KJ)	-16161	-16155	342
Sum of energy loss (KJ)	16218	16221	15560	
Abs ICE mechanical energy-sum of energy loss - extra energy stored in battery  (KJ)	57	66	25	

## Chapter 9

# Conclusions, Future Work and Publications

## 9.1 Conclusions

This dissertation has fueled the research on powertrain electrification in two distinct directions, i.e., the pursuit for a state-of-the-art real-time implementable near-optimal control strategy for the energy management system of a multi-mode electrified powertrain and buttressing the pursuit for a novel multi-mode powertrain configurations with the development of an online simulation platform for rapidly evaluating their performance. The dissertation starts with delineating the fundamentals of powertrain electrification and the energy management system, followed by highlighting the existing literature on energy management strategies and elaborating the implementation of a well-accredited energy management strategy. The dissertation culminates in finding that applying an *Asynchronous actor-critic* agent would be the future of real-time optimal control of energy management systems.

It has been corroborated in online simulation platform that if properly articulated, an *Asynchronous advantage actor-critic*-based agent can overwhelm the well-accredited *equivalent consumption minimization strategy* in real-world driving conditions with no prior knowledge of the drive cycle ahead. Although pre-collected driving data were employed to train the actor-critic agent in the online simulation platform, the proposed framework demonstrates that the agent’s “on-the-fly” training is possible with just an onboard multi-core CPU and a Markov chain model. The training time is inversely proportional to the number of CPU core used for training. The demonstration leaves a perfect avenue for further study on multiple aspects of the actor-critic agent, i.e., a parametric study and sensitivity analysis on various hyper-parameters, a few of which have been considered in this dissertation, fine-tuning the implementation before its software-in-the-loop and subsequently hardware-in-the-loop validation.

The dissertation meticulously documented the differences of reinforcement learning implementations between simple power-split and multi-mode electrified powertrains and prescribed an ingenious method of conflation between reinforcement learning and equivalent consumption minimization strategy for circumventing the hindrances posed by a multi-mode electrified powertrain.

However, there are a few aspects, such as reducing emission while minimizing fuel consumption, which has not been considered in the current research. Hence, the current EMS cannot be deployed in a real HEV immediately. To deploy the DRL as an EMS policy for a real HEV, its immediate cost function must include a term penalizing the tailpipe emission. Moreover, a naïve DRL cannot be deployed in a real HEV, although it is inevitable that the DRL’s performance will improve over time, and such constraint is conducive for HEV’s safe operation. It is customary to train the DRL encompassing rudimentary objectives on standard federal drive cycles before deploying on the road. Once the DRL is on the road, it will keep on learning from the most contemporary drive cycles and make itself near-optimal for any unknown drive cycles belonging to the same contemporary driving trend.

While articulating a near-optimal energy management strategy for a multi-mode electrified powertrain, an urge to model the powertrain dynamics during mode shift events has been sensed. Such an urge fueled the research towards the second contribution of this dissertation, i.e., articulating a low-frequency coordinated control for grappling torque fluctuation at the transmission output during mode shift events. It is noteworthy to mention that the investigation has also been ignited from a previous research on “*Search for novel configurations of multi-mode powertrains*”. The selection of a few novel powertrain configurations through a screening process of copious

candidate accommodates a performance evaluation step, exactly where the necessity of high-fidelity powertrain model knocks the door. It has been corroborated here that a coordinated control is strictly necessary for a high-fidelity powertrain model to emulate a real physical powertrain. Following the same path as its predecessor, this contribution also opens new avenue for various directions such as impact of high-frequency coordinated control on mode scheduling strategy, detailed clutch dynamics during all types of mode shift events, and impact of high-fidelity control oriented component models on the coordinated control.

## 9.2 Future Work

The dissertation reveals only the tip of an ice-berg. A multitude of commendable research topics on powertrain electrification are awaiting to be ignited as a derivative of this dissertation. Those research topics are enumerated as follows:

1. The mobility of the future is going to be more and more smart and sustainable. EMS has already been coupled with the connected vehicle technology. EMS's responsibility is to optimize the production and distribution of on-board generated power for a given drive cycle. EMS does not have any role in optimizing the vehicle's speed ahead. Connected vehicle technology helps to optimize the vehicle's speed for the short-term future in real-time. Hence, the EMS augmented with connected vehicle technology will provide more prudent solution to the overall energy management problem in real-time.
2. As described in the 3<sup>rd</sup> chapter, the development phases of a utilitarian energy management system starts with the mathematical conceptualization and ends

with on-road validation. The dissertation starts with the mathematical conceptualization but ends with only model-in-the-loop simulation. Validation of the energy management strategies articulated in various chapters need to be validated in hardware-in-the-loop simulation followed by dynamometer test. The DRL strategy is expected to work perfectly if the DRL does not need to be trained online in the HIL. A trained DRL from the model-in-the-loop simulation can be imported into the HIL simulation and tested for its adaptability. The strategy needs to be implemented through some modification if it needs to be trained online in a HIL bench.

3. To materialize the strategies on an embedded system from a model-in-the-loop platform several apposite changes must be implemented in the real-time optimization and in the plant model. High-frequency controller must be designed for high-fidelity control-oriented plant model.
4. The internal combustion engine is not going to be superseded by electric motors any sooner than 20 years down the line. Hence, hybrid electric vehicles requires more attention from scholars. Advanced hybrid-specific engine technologies and controls must be incorporated in the future research plan. The first step towards that would be development of high-fidelity control oriented models of internal combustion engines and validation of those models with engine-dynamometer data.
5. The literature is inundated with copious number of model-in-the-loop controllers' implementation, whose validity is often questioned by other scholars. Chassis dynamometer test of a certain application vehicle with a certain energy

management strategy and powertrain configuration will facilitate the scholars with the opportunity of validating their online simulation-based controllers. Such test results will also assist the scholars to improve their vehicle plant model.

6. Since, there is no exclusive rule-of-thumb for selecting the number of hidden layers and neurons per hidden layer for DRL, a sensitivity analysis would be ideal to analyze the effect of DNN's architecture on the convergence performance of the DRL.
7. Since most OEMs are moving towards plug-in HEV rather than just HEV, future work should be focused on articulating EMS to minimize fuel consumption for the plug-in HEVs. Plug-in HEVs leverage the bigger battery capacity and external charging option, and hence the EMS of a plug-in HEV can let the battery deplete up to the lower limit without satisfying the charge sustainability criteria. However, the EMS of a plug-in HEV should not let the battery deplete below the lower limit until the next charging station. Hence, the distance from the next charging station is a piece of important state information required for articulating the DRL-based EMS.
8. In the current research, A3C is prioritized over A2C, i.e., synchronous advantage actor-critic, only due to its convergence rapidity. However, literature corroborates the necessity of more computational power for the implementation of A3C compared to A2C. Scholars are well acquainted with the fact that onboard microprocessors always face the dearth of computational power, and the automotive industry always tries to optimize the balance between computational



power and performance. Looking from the actual vehicular implementation’s perspective, a comparative analysis on the required computational power among different DRL policies would be noteworthy in the future.

9. As far as the energy management system of a real HEV in the real-world driving scenario is concerned, it would be unfair to denote it as an ideal Markov decision process (MDP) since there could be many hidden states of the MDP which might or might not conform with the Markovian property, i.e., the immediate next state might or might not be related only to the current state. In non-MDP or partially observable MDP, the agent can get a different next state or reward if the same action is applied in the same state but in a different time-step. Another challenge in real-world driving can engender from the inter-relation between states from different time-steps. The dependency of the current state on a previous state makes the job difficult for the DRL. A recurrent neural network such as long-short term memory becomes a wise option for the network in such a case.
10. All the collected, generated, and testing drive cycles are real-world drive cycles. However, they are not explicitly labeled as aggressive drive cycles. Future work would be focused on testing the DRL’s performance in a standard aggressive drive cycle such as US06.
11. The training of a DRL-based policy employed in the EMS of a specific HEV can be time-consuming. If the designer intends to articulate a DRL-based EMS for another HEV with different specifications and powertrain architecture, would the designer have to start training the new DRL from scratch? The answer

is no. With the advent of *Transfer learning*, the learned parameters from the former HEV's DRL can be imported to the newer HEV's DRL swiftly, and the latest learning process can be expedited significantly. The efficacy of *Transfer learning* will be expounded in future work.

## 9.3 Publications

### 9.3.1 Published Journal papers

1. **A. Biswas** and A. Emadi, "Energy management systems for electrified powertrains: State-of-the-art review and future trends," in IEEE Transactions on Vehicular Technology, vol. 68, no. 7, pp. 6453-6467, July 2019.
2. **A. Biswas**, P. G. Anselma, A. Rathore, and A. Emadi, "Effect of Coordinated Control on Real-time Optimal Mode Selection for Multi-mode Hybrid Electric Powertrain", Accepted in Applied Energy.
3. P. G. Anselma, **A. Biswas**, G. Belingardi, and A. Emadi, "Rapid assessment of the fuel economy capability of parallel and series-parallel hybrid electric vehicles," in Applied Energy, vol. 275, p. 115319, 2020.
4. R. Ostadian, J. Ramoul, **A. Biswas**, and A. Emadi, "Intelligent Energy Management Systems for Electrified Vehicles: Current Status, Challenges, and Emerging Trends," in IEEE Open Journal of Vehicular Technology, vol. 1, pp. 279-295, July 2020.

### 9.3.2 Journal papers under review

1. **A. Biswas**, and A. Emadi, “Application of Asynchronous Actor-Critic Agent in Model-in-the-loop Simulation of Energy Management System for an Electrified Powertrain”, submitted to Applied Energy.

### 9.3.3 Published Conference papers

1. **A. Biswas**, R. Gu, P. Kollmeyer, R. Ahmed, and A. Emadi “Simultaneous state and parameter estimation of li-ion battery with one state hysteresis model using augmented unscented kalman filter,” 2018 IEEE Transportation Electrification Conference and Expo (ITEC), Long Beach, CA, 2018, pp. 1065-1070.
2. **A. Biswas** P. G. Anselma, and A. Emadi, “Real-time optimal energy management of electrified powertrains with reinforcement learning,” 2019 IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, USA, 2019, pp. 1-6.
3. **A. Biswas**, P. G. Anselma, A. Rathore, and A. Emadi, “Comparison of Three Real-Time Implementable Energy Management Strategies for Multi-mode Electrified Powertrain,” 2020 IEEE Transportation Electrification Conference and Expo (ITEC), Chicago, IL, USA, 2020, pp. 514-519.
4. P. G. Anselma, **A. Biswas**, J. Roeleveld, G. Belingardi, and A. Emadi, “Multi-fidelity near-optimal on-line control of a parallel hybrid electric vehicle powertrain,” 2019 IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, USA, 2019, pp. 1-6.

5. C. Vidal, M. Haußmann, D. Barroso, P. M. Shamsabadi, **A. Biswas**, E. Chemali, R. Ahmed, and A. Emadi,, "Hybrid Energy Storage System State-Of-Charge Estimation Using Artificial Neural Network For Micro-Hybrid Applications," 2018 IEEE Transportation Electrification Conference and Expo (ITEC), Long Beach, CA, 2018, pp. 1075-1081.
6. Anselma, P., **Biswas, A.**, Bruck, L., Amirfarhangi Bonab, S. et al., "Accelerated Sizing of a Power Split Electrified Powertrain," SAE Int. J. Adv. and Curr. Prac. in Mobility 2(5):2701-2711, 2020.
7. Bruck, L., Lempert, A., Amirfarhangi Bonab, S., Lempert, J., **Biswas, A.** et al. "A Dynamic Programming Algorithm for HEV Powertrains Using Battery Power as State Variable," SAE Technical Paper 2020-01-0271, 2020.
8. Bruck, L., Amirfarhangi Bonab, S., Lempert, A., **Biswas, A.** et al., "An Iterative Histogram-Based Optimization of Calibration Tables in a Powertrain Controller," SAE Technical Paper 2020-01-0266, 2020.

# Appendices

## Appendix A

Efficiency and fuel contour maps of a few ICE models, SOC-OCV and efficiency contour map of the battery used in this dissertation

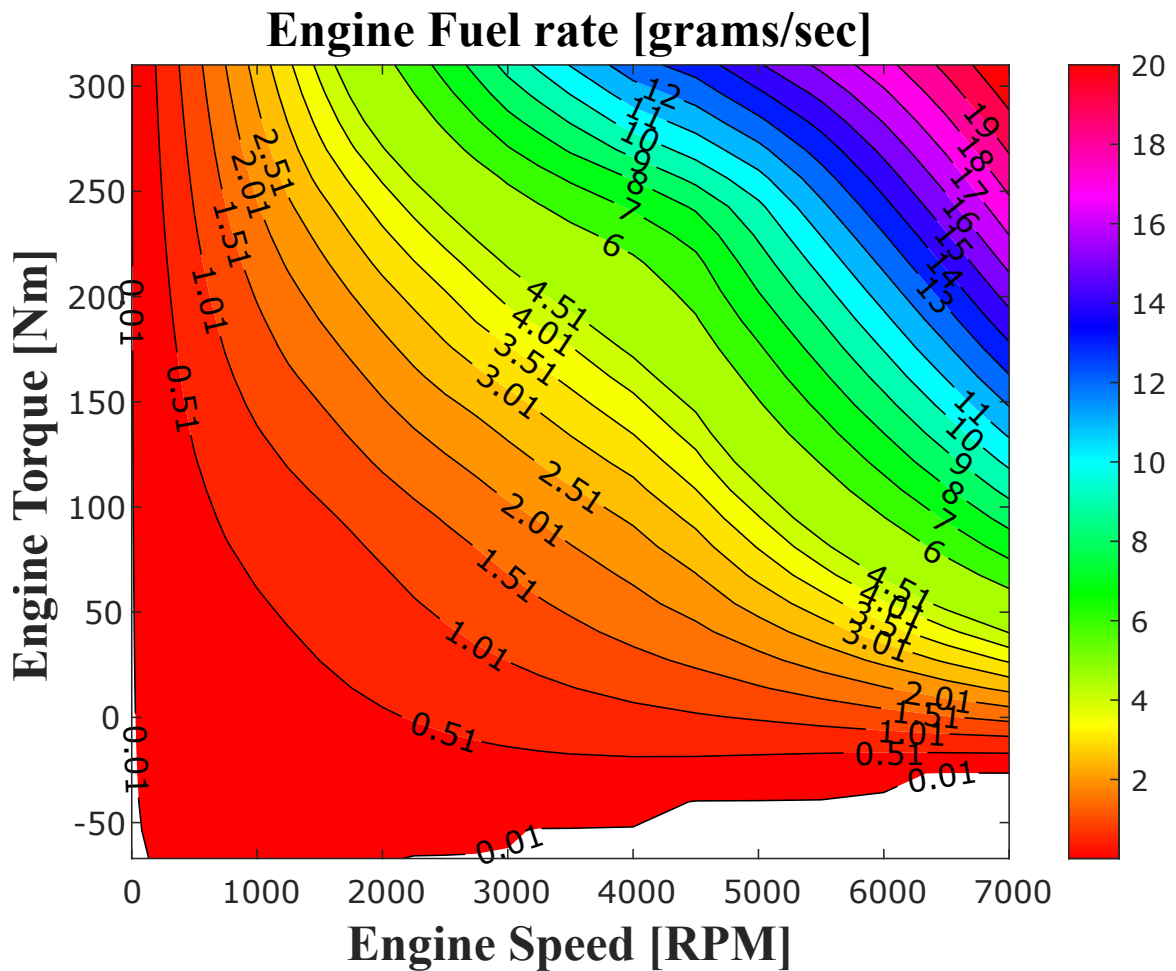


Figure A.1: Fuel map of a 2 Liter spark-ignition ICE used in the multi-mode electrified powertrain’s simulation in this dissertation. The ICE’s data is obtained from U.S. Environmental Protection Agency’s (EPA’s) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. This ICE data have been used for the articulation of reinforcement learning agent’s simulation.

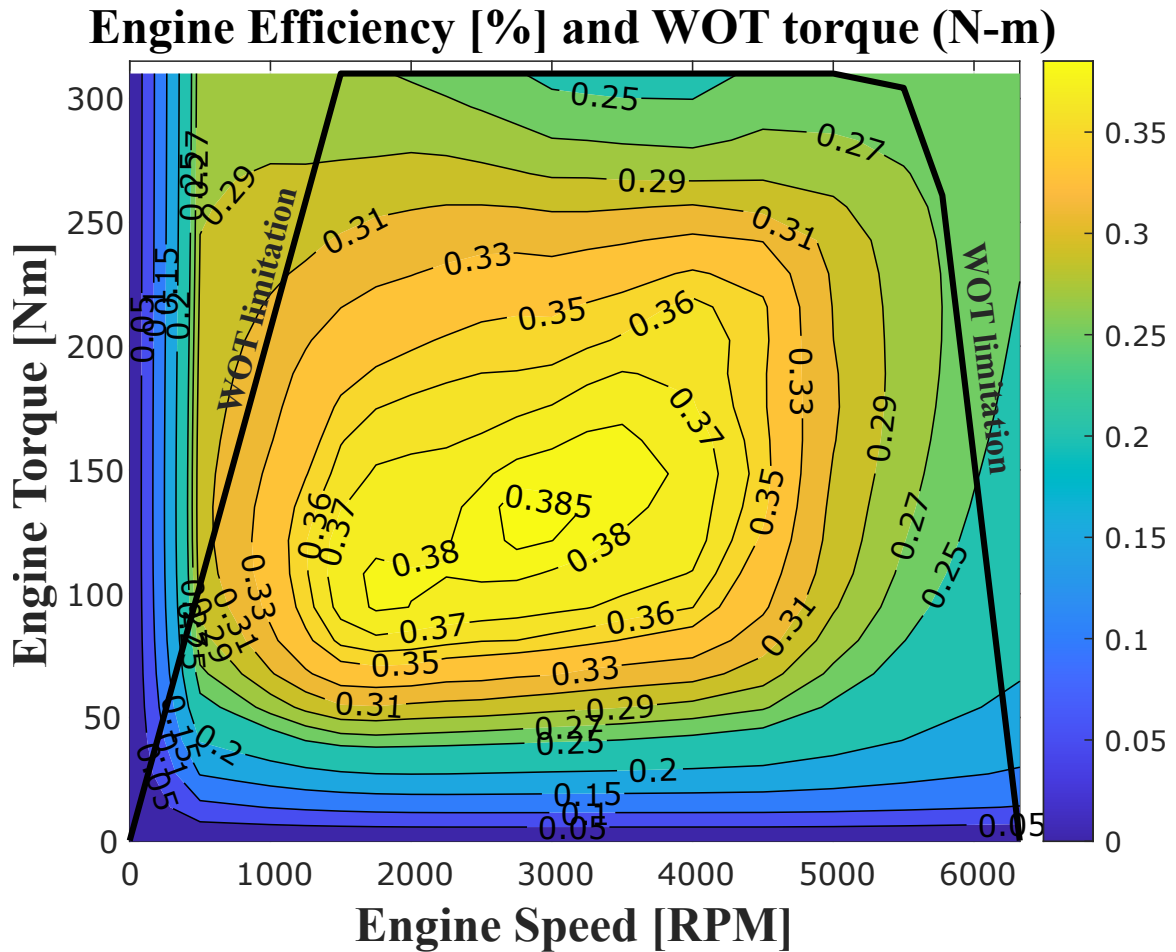


Figure A.2: Efficiency map of a 2 Liter spark-ignition ICE used in the multi-mode electrified powertrain’s simulation in this dissertation. The ICE’s data is obtained from U.S. Environmental Protection Agency’s (EPA’s) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202]. This ICE data have been used for the articulation of reinforcement learning agent’s simulation.



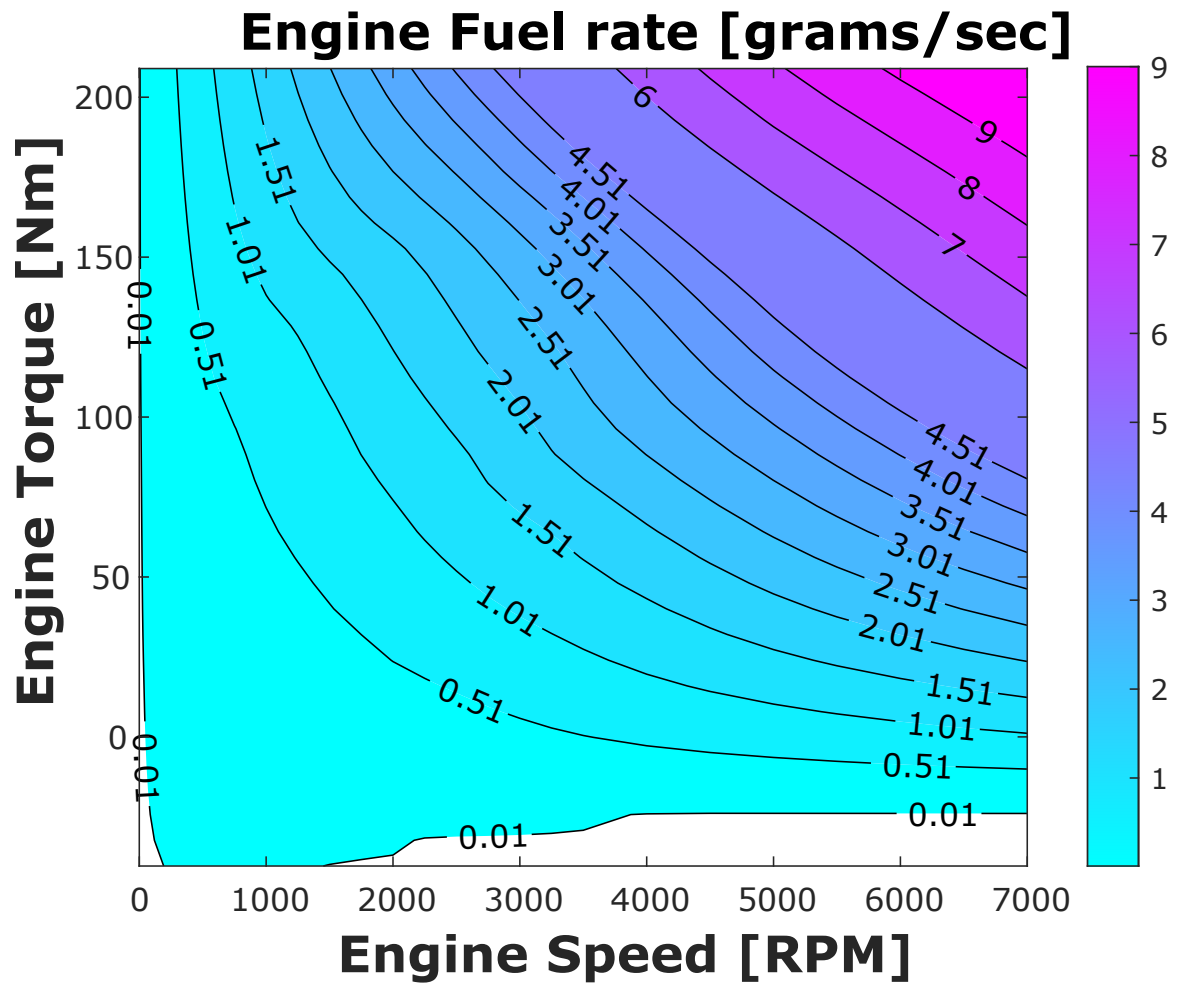


Figure A.3: Some extra Fuel map of another 2 Liter spark-ignition ICE. The ICE's data is obtained from U.S. Environmental Protection Agency's (EPA's) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202].

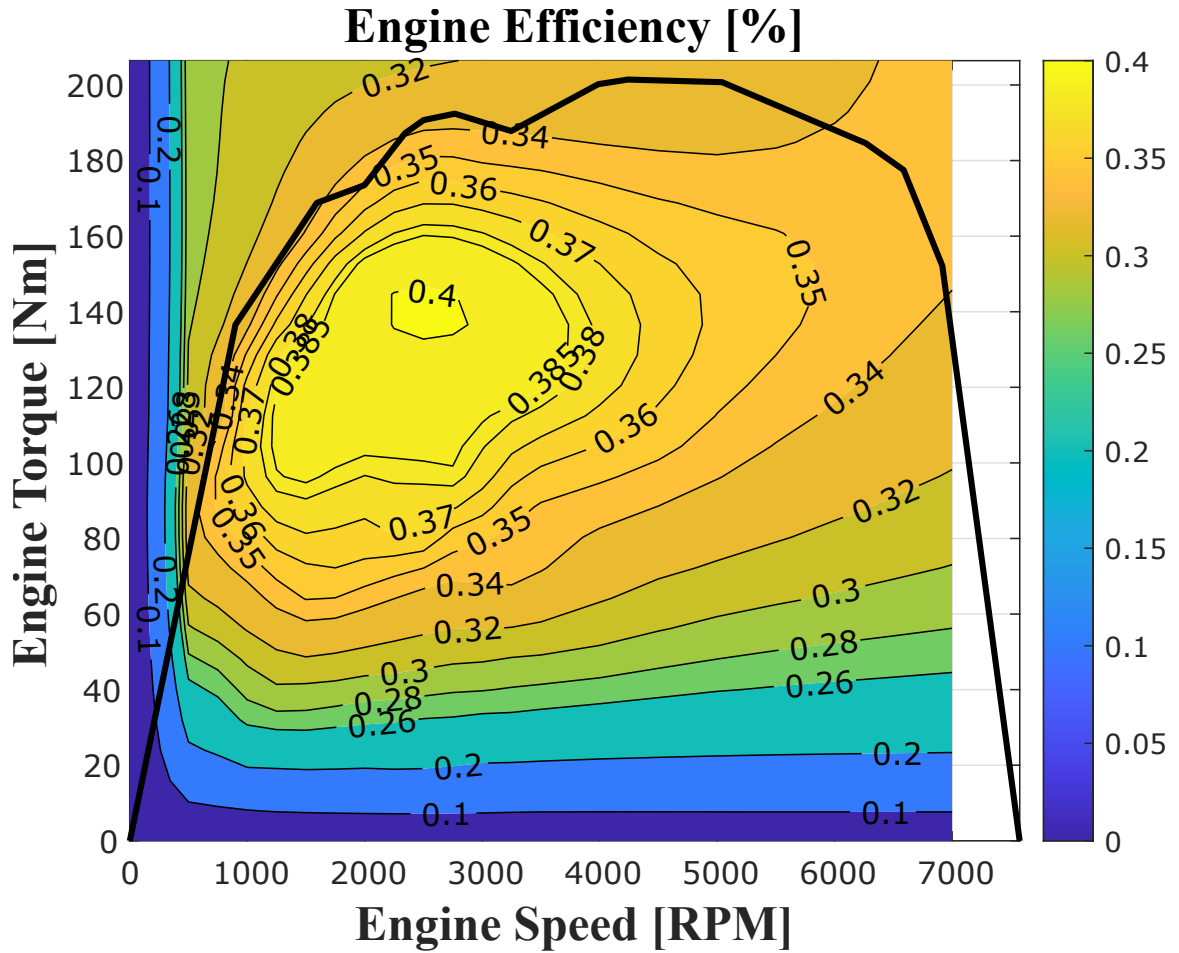


Figure A.4: Some extra efficiency map of another 2 Liter spark-ignition ICE. The ICE's data is obtained from U.S. Environmental Protection Agency's (EPA's) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202].

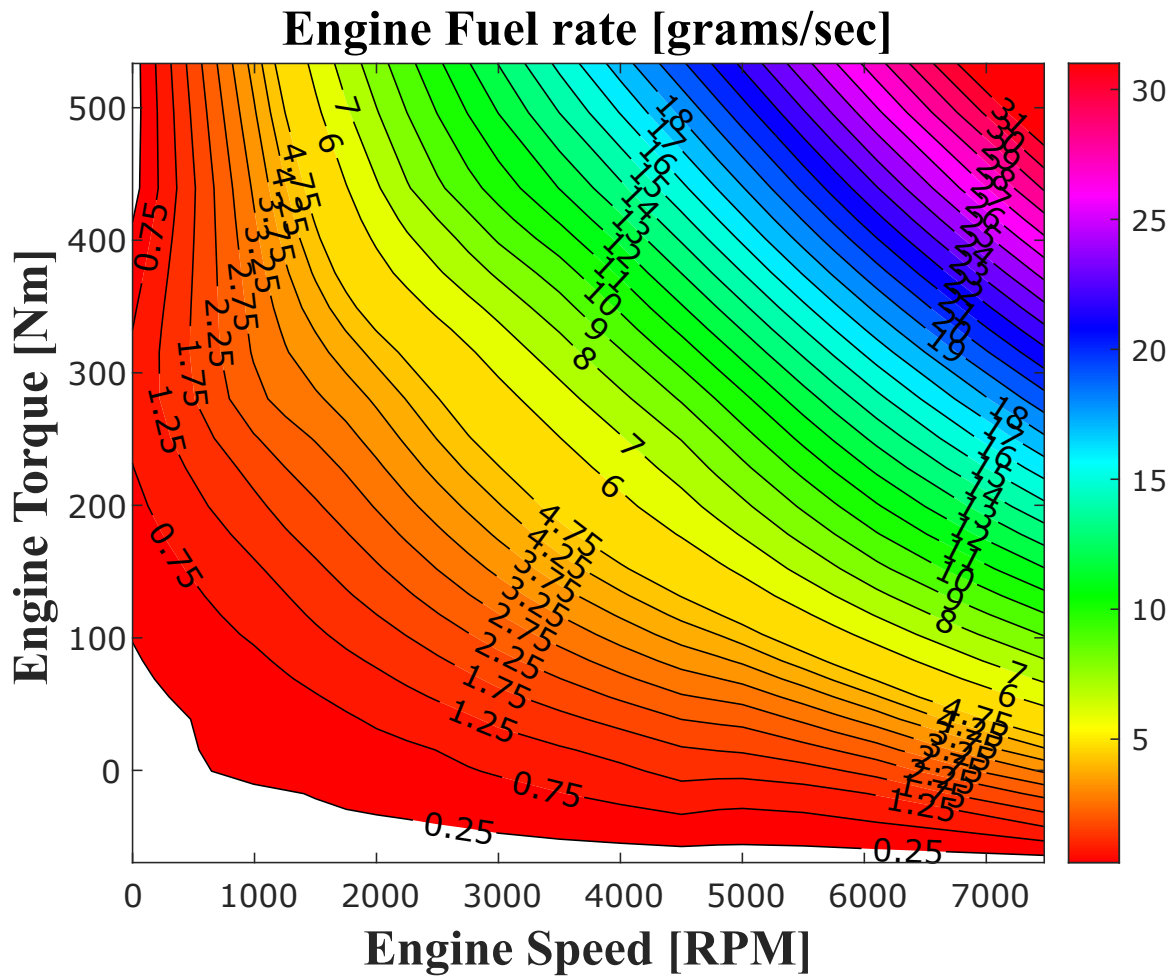


Figure A.5: Some extra Fuel map of another 2.7 Liter spark-ignition ICE. The ICE's data is obtained from U.S. Environmental Protection Agency's (EPA's) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202].

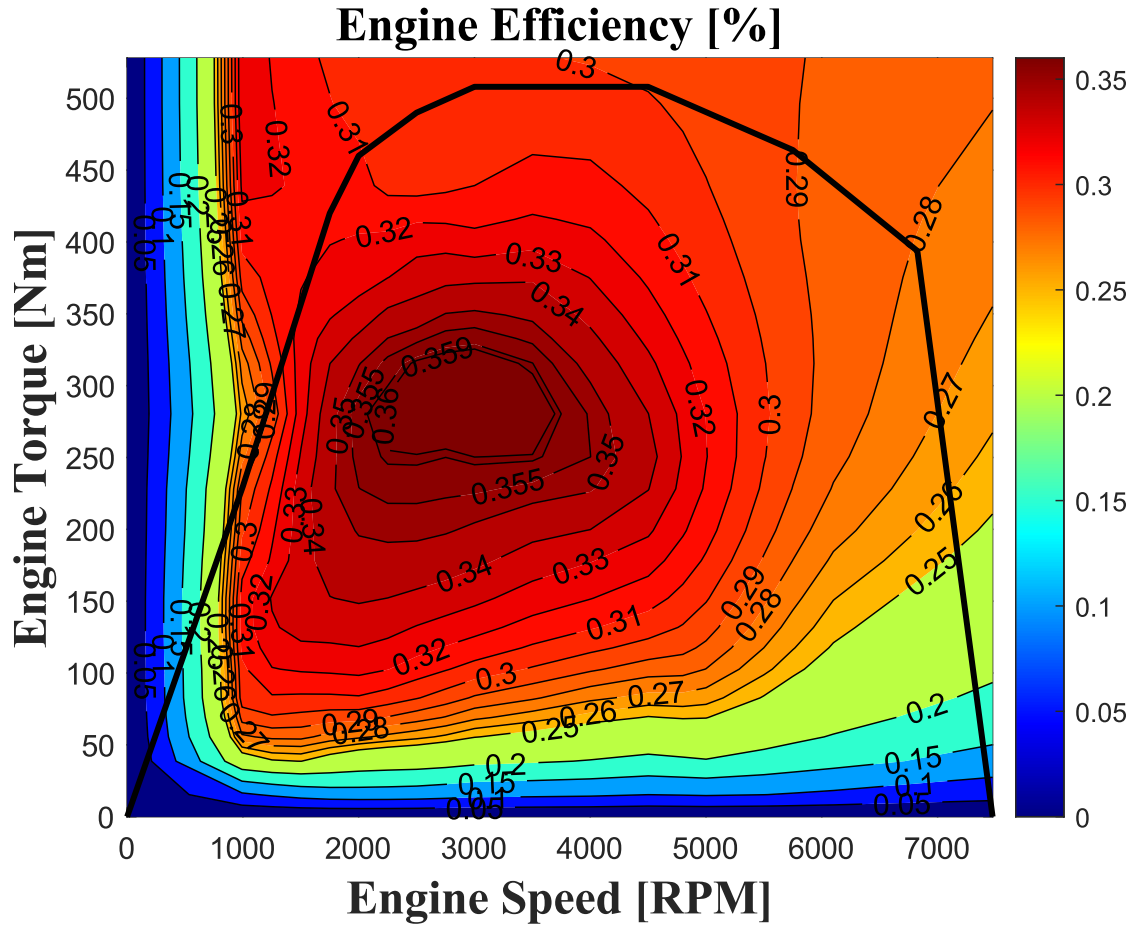


Figure A.6: Some extra efficiency map of another 2.7 Liter spark-ignition ICE. The ICE's data is obtained from U.S. Environmental Protection Agency's (EPA's) Advanced Light-Duty Powertrain and Hybrid Analysis (ALPHA) tool [202].

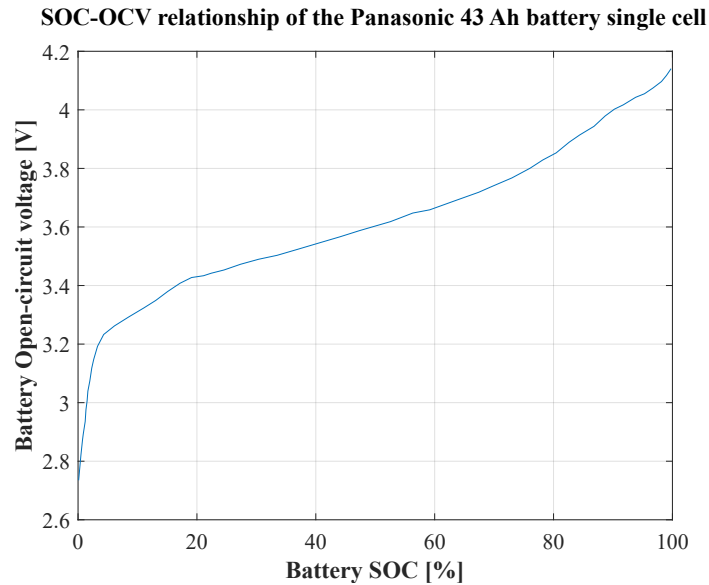


Figure A.7: SOC-OCV relationship of a Panasonic battery cell used in the multi-mode electrified vehicle used for online simulation.

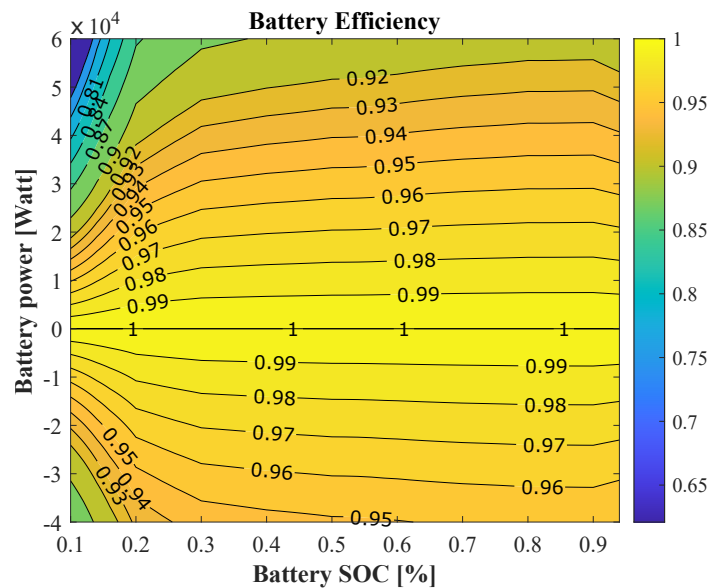


Figure A.8: Efficiency of a Panasonic battery cell used in the multi-mode electrified vehicle used for online simulation.

## Appendix B

Detailed Pseudo-code for  
Q-learning algorithm used for  
offline simulation in this  
dissertation

---

**Algorithm 4:** Detailed Pseudo-code for Q-learning based training of a reinforcement learning agent

---

**Begin**

- 1) **Vehicle Parameters Initialization** : *Road load co –*  
*efficients, vehicle<sub>mass</sub>, vehicle<sub>inertia</sub>, wheel<sub>radius</sub>, Gear ratio<sub>final drive</sub>*  
*Air density ( $\rho_{air}$ ), vehicle<sub>height</sub>, vehicle<sub>width</sub>, Aero drag coefficient*
  - 2) **Load Drive Cycles**
  - 3) **Power-demand calculation for each of the above drive cycles:**  
 $Power_{Demand} =$   
*powerDmndCals (vel, acc, grade, vehicle<sub>mass</sub>, vehicle<sub>inertia</sub>, wheel<sub>radius</sub>*  
*, Road load co – efficients, FDR, height, width, aeroDragCoeff)*
  - 4) **Discretizing power demand (1<sup>st</sup> state variable) in nonlinear way to discretize based on its histogram distribution:**  
 $State_{power\_index} = nonlinspace(Power_{Demand})$   
 $State_{power\_index} \in \{State_{power_1}, State_{power_1}, \dots, State_{power_N}\}$
  - 4) **Discretizing SOC (2<sup>nd</sup> state variable) in a linear way:**  
 $State_{SOC\_index} \in \{0.1 : (0.8/40) : 0.9\}$
  - 5) **Initializing state vector. It is a cell vector having a size of**  
 $(Size_{SOC\_index} \times Size_{power\_index}, 1):$   
 $Size_{state} = length(Size_{SOC\_index}) \times length(Size_{power\_index})$   
 $State = cell(Size_{state}, 1)$
  - 6) **Completion of State vector:** for  $ii = 1 : 1 : State_{SOC\_index}$  **do**  
     for  $jj = 1 : 1 : State_{power\_index}$  **do**  
          $position = (ii - 1) * length(State_{power\_index}) + jj;$   
          $State\{position, 1\} = [State_{SOC\_index}(ii), State_{power\_index}(jj)];$   
          $State\{position, 2\} = single([ii, jj]);$   
     **end**  
**end**
  - 7) **Discretization of Action variables and Initialization of Action vector:** Choose the appropriate action variables and discretize them uniformly. Initialize Action vector by following step 5).  
 While articulating Action vector, feasibility of two or more action variable's combination should be considered judiciously.
  - 8) **Construction of Immediate Cost matrix (very important matrix):**  
 $Size_{Immediate\ Cost\ Matrix} = length(position) \times length(Action\ vector)$   
 $Immediate\ Cost\ Matrix = inf \times ones(position, length(Action\ vector));$
-

---



---

```

for  $mm = 1 : 1 : position$  do
  for  $nn = 1 : 1 : length(Action\ vector)$  do
    Immediate Cost Matrix( $mm, nn$ ) =
      Immediatecost( $state_{mm}, action_{nn}, other\ inputs$ );
    , where Immediatecost is a function defining the immediate cost
      corresponding to certain combination of action index and state
      index.

```

**9) Q-learning: An off-policy TD control algorithm (Refer to Reinforcement Learning: An Introduction):**

9.a) *Initialize the Q-matrix:*

$q = randi(500, size(Immediate\ Cost\ Matrix));$

```

for  $mm = 1 : 1 : position$  do

```

```

  for  $nn = 1 : 1 : length(Action\ vector)$  do

```

```

    if Immediate Cost Matrix( $mm, nn$ ) ==  $\infty$  then

```

```

       $q(mm, nn) = \infty;$ 

```

9.b) *Set Discount Factor:*  $\gamma = 0.9$

9.c) *Set Learning Rate:*  $\alpha = 0.2$

9.d) *Set maximum number of iterations:*  $Max.Iteration = 2000$

9.e) *Set Initial value of Battery SOC:*  $SOC_{initial} = 50\%$

**Main section of Q-learning:**  $CS = currentstate$   $NS = nextstate$

---



---



---

```

for  $i=1:Max.Iteration$  do
   $SOC_{CS} = SOC_{initial}$ ; Calculate  $Power_{Demand_{CS}}$ 
  for  $t = 1 : 1 : length(drive\ cycle)$  do
    1) With maximum likelihood find the index of present SOC and
       power demand. Find the  $State_{index}$  of  $CS$  from the state vector.
    2)  $CS = (indx_{SOC_{CS}} - 1) * length(State_{power_{index}}) + indx_{power_{CS}}$ ; 3)
       Calculate OCV, internal resistance of the battery at the SOC of
        $CS$ .
    4) Selection of action for the current state (CS)
        $temp = random\ number\ between\ 0\ to\ 1\ (temp \in [0, 1])$ 
       Implementing  $\epsilon$ -greedy policy
       if  $temp < 0.1$  then
         Action will be selected randomly from the  $Vector_{Action}$ 
       else
         Action will be selected by greedy policy
          $Action_{CS} = \arg\ min_{Action_{kk}} \|(q(CS, :))\|$ 
       5) Control variables for the current state are selected based on
           $Action_{CS}$ 
       6) Once, the control variables are selected, next state value can be
          calculated using the state dynamics equation.
       7) Calculate the next SOC value based on the battery power
          demand:
           $SOC_{NS} = SOC_{CS} -$ 
           $ddtSOC(Batt_{Pow}, OCV, Intrnl\ Resi_{.batt}, Max.\ Capa_{.batt})$ ;
          Calculate  $Power_{Demand_{NS}}$  From pre-calculated values of  $Power_{demand}$ 
       8) With maximum likelihood find the index of present SOC and
          power demand. Find the  $State_{index}$  of  $NS$  from the state vector.
       9)  $NS = (indx_{SOC_{NS}} - 1) * length(State_{power_{index}}) + indx_{power_{NS}}$ ;
       10) Selection of action for the next state (NS): Same approach will
          be followed as step (4).
       11) Update q-values as per Bellman's equation:(Important step)
           $q(CS, Action_{CS}) =$ 
           $q(CS, Action_{CS}) + \alpha * (Immediate\ Cost\ Matrix(CS, Action_{CS}) +$ 
           $\gamma \times q_{min} - q(CS, Action_{CS}))$ ;
       Check whether the episode has completed
       Set current state as next state:
           $Power_{Demand_{CS}} = Power_{Demand_{NS}}$ ;  $SOC_{CS} = SOC_{NS}$ ;

```

---

# References

- [1] Sales of new vehicles 2005-2018. [Accessed: 2019-12-30]. [Online]. Available: <http://www.oica.net/category/sales-statistics/>
- [2] A. Emadi, *Advanced electric drive vehicles*. Boca Raton, FL: CRC Press, Oct. 2014.
- [3] B. Bilgin, P. Magne, P. Malysz, Y. Yang, V. Pantelic, M. Preindl, A. Korobkine, W. Jiang, M. Lawford, and A. Emadi, “Making the case for electrified transportation,” *IEEE Transactions on Transportation Electrification*, vol. 1, no. 1, pp. 4–17, June 2015.
- [4] A. Biswas and A. Emadi, “Energy management systems for electrified powertrains: State-of-the-art review and future trends,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 6453–6467, July 2019.
- [5] Hybrid electric vehicle sales in u.s. by model, 1999 to 2017. [Accessed: 2020-01-05]. [Online]. Available: <https://afdc.energy.gov/data/10301>
- [6] T. M. Grewe, B. M. Conlon, and A. G. Holmes, “Defining the general motors 2-mode hybrid transmission,” in *SAE Technical Paper*. SAE International, 04 2007. [Online]. Available: <https://doi.org/10.4271/2007-01-0273>

- [7] J. Kim, T. Kim, B. Min, S. Hwang, and H. Kim, “Mode control strategy for a two-mode hybrid electric vehicle using electrically variable transmission (evt) and fixed-gear mode,” *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 793–803, March 2011.
- [8] S. L. Poczter and L. M. Jankovic, “General motors: The electric revolution,” *Journal of Business Case Studies (JBKS)*, vol. 10, no. 2, pp. 185–190, 2014.
- [9] R. Ostadian, J. Ramoul, A. Biswas, and A. Emadi, “Intelligent energy management systems for electrified vehicles: Current status, challenges, and emerging trends,” *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 279–295, 2020.
- [10] C.-C. Lin, H. Peng, and J. W. Grizzle, “A stochastic control strategy for hybrid electric vehicles,” in *Proc. 2004 Am. Control. Conf.*, vol. 5, Jun. 2004, pp. 4710–4715 vol.5.
- [11] L. Johannesson, M. Asbogard, and B. Egardt, “Assessing the potential of predictive control for hybrid vehicle powertrains using stochastic dynamic programming,” *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 1, pp. 71–83, Mar. 2007.
- [12] S. D. Cairano, D. Bernardini, A. Bemporad, and I. V. Kolmanovsky, “Stochastic mpc with learning for driver-predictive vehicle control and its application to hev energy management,” *IEEE Transactions on Control Systems Technology*, vol. 22, no. 3, pp. 1018–1031, May 2014.
- [13] H. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, I. V. Kolmanovsky, and S. Di Cairano, “Mpc-based energy management of a power-split hybrid electric

- vehicle,” *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 593–603, May 2012.
- [14] S. Zhang, R. Xiong, and F. Sun, “Model predictive control for power management in a plug-in hybrid electric vehicle with a hybrid energy storage system,” *Applied Energy*, vol. 185, pp. 1654 – 1662, 2017, clean, Efficient and Affordable Energy for a Sustainable Future. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261915016086>
- [15] G. Paganelli, S. Delprat, T. M. Guerra, J. Rimaux, and J. J. Santin, “Equivalent consumption minimization strategy for parallel hybrid powertrains,” in *IEEE 55th Veh. Technol. Conf.*, vol. 4, May 2002, pp. 2076–2081 vol.4.
- [16] L. Serrao and G. Rizzoni, “Optimal control of power split for a hybrid electric refuse vehicle,” in *Proc. 2008 Am. Control Conf.*, June 2008, pp. 4498–4503.
- [17] N. Kim, S. Cha, and H. Peng, “Optimal control of hybrid electric vehicles based on Pontryagin’s minimum principle,” vol. 19, no. 5, pp. 1279–1287, Sep. 2011, Accessed: 2019-03-12. [Online]. Available: DOI:10.1109/TCST.2010.2061232
- [18] S. Ebbesen, P. Elbert, and L. Guzzella, “Battery state-of-health perceptive energy management for hybrid electric vehicles,” vol. 61, no. 7, pp. 2893–2900, Sep. 2012, Accessed: 2017-05-28. [Online]. Available: DOI: 10.1109/TVT.2012.2203836
- [19] C. Musardo, G. Rizzoni, Y. Guezennec, and B. Staccia, “A-ECMS: An adaptive algorithm for hybrid electric vehicle energy management,” *European*

- Journal of Control*, vol. 11, no. 4, pp. 509 – 524, 2005, Accessed: 2018-04-05. [Online]. Available: DOI:10.3166/ejc.11.509-524
- [20] P. Pisu and G. Rizzoni, “A comparative study of supervisory control strategies for hybrid electric vehicles,” vol. 15, no. 3, pp. 506–518, May 2007, Accessed: 2017-11-14. [Online]. Available: DOI:10.1109/TCST.2007.894649
- [21] J. T. B. A. Kessels, M. W. T. Koot, P. P. J. van den Bosch, and D. B. Kok, “Online energy management for hybrid electric vehicles,” vol. 57, no. 6, pp. 3428–3440, Nov 2008, Accessed: 2018-09-24. [Online]. Available: DOI:10.1109/TVT.2008.919988
- [22] A. Chasse and A. Sciarretta, “Supervisory control of hybrid powertrains: An experimental benchmark of offline optimization and online energy management,” *Control Eng. Pract.*, vol. 19, no. 11, pp. 1253 – 1265, 2011. [Online]. Available: DOI:10.1016/j.conengprac.2011.04.008
- [23] F. Tianheng, Y. Lin, G. Qing, H. Yanqing, Y. Ting, and Y. Bin, “A supervisory control strategy for plug-in hybrid electric vehicles based on energy demand prediction and route preview,” vol. 64, no. 5, pp. 1691–1700, May 2015, Accessed: 2019-02-15. [Online]. Available: DOI:10.1109/TVT.2014.2336378
- [24] S. Yang, W. Wang, F. Zhang, Y. Hu, and J. Xi, “Driving-style-oriented adaptive equivalent consumption minimization strategies for HEVs,” vol. 67, no. 10, pp. 9249–9261, Oct 2018, Accessed: 2019-03-07. [Online]. Available: DOI:10.1109/TVT.2018.2855146

- [25] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, 1st ed. Cambridge, MA, USA: MIT Press, 1998.
- [26] V. Mnih *et al.*, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.
- [27] X. Lin, Y. Wang, P. Bogdan, N. Chang, and M. Pedram, “Reinforcement learning based power management for hybrid electric vehicles,” in *Proc. of 2014 IEEE/ACM Int. Conf. Computer-Aided Design*, ser. ICCAD ’14. Piscataway, NJ, USA: IEEE Press, 2014, pp. 32–38.
- [28] The design, manufacture and sales of automobiles drive economic growth around the world. [Accessed: 2019-12-30]. [Online]. Available: <http://www.oica.net/category/economic-contributions/>
- [29] U.s. epa’s inventory of u.s. greenhouse gas emissions and sinks: 1990-2017. [Accessed: 2019-12-30]. [Online]. Available: <http://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks>
- [30] Greenhouse gas emissions by economic sector, canada, 1990 to 2017. [Accessed: 2019-12-31]. [Online]. Available: <https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/greenhouse-gas-emissions.html>
- [31] M. Porter and C. Van der Linde, “Green and competitive: ending the stalemate,” *The Dynamics of the eco-efficient economy: environmental regulation and competitive advantage*, vol. 33, 1995.

- [32] S. C. Debnath, “Environmental regulations become restriction or a cause for innovation – a case study of toyota prius and nissan leaf,” *Procedia - Social and Behavioral Sciences*, vol. 195, pp. 324 – 333, 2015, world Conference on Technology, Innovation and Entrepreneurship.
- [33] E. Chemali, M. Preindl, P. Malysz, and A. Emadi, “Electrochemical and electrostatic energy storage and management systems for electric drive vehicles: State-of-the-art review and future trends,” *IEEE Trans. Emerg. Sel. Topics Power Electron.*, vol. 4, no. 3, pp. 1117–1134, Sep. 2016.
- [34] G. Wu, X. Zhang, and Z. Dong, “Powertrain architectures of electrified vehicles: Review, classification and comparison,” *J. Franklin Inst.*, vol. 352, no. 2, pp. 425 – 448, 2015.
- [35] P. Suntharalingam, A. Emadi, and M. Zhang, “Hybrid electric powertrains: Current status, future trends, and electro-mechanical integration methods,” *Int. J. Powertrains*, vol. 3, no. 3, pp. 319–349, 2014.
- [36] A. Emadi, *Handbook of automotive power electronics and motor drives*. Boca Raton, FL: CRC press, May 2005.
- [37] Y. Yang, K. Arshad-Ali, J. Roeleveld, and A. Emadi, “State-of-the-art electrified powertrains-hybrid, plug-in, and electric vehicles,” *Int. J. Powertrains*, vol. 5, no. 1, pp. 1–29, 2016.
- [38] J. Liu and H. Peng, “A systematic design approach for two planetary gear split hybrid vehicles,” *Vehicle Syst. Dyn.*, vol. 48, no. 11, pp. 1395–1412, 2010.

- [39] P. G. Anselma, Y. Huo, J. Roeleveld, A. Emadi, and G. Belingardi, “Rapid optimal design of a multimode power split hybrid electric vehicle transmission,” *Proc. Inst. Mech. Eng. D*, vol. 233, no. 3, pp. 740–762, 2019.
- [40] M. A. Miller, A. G. Holmes, B. M. Conlon, and P. J. Savagian, “The gm “voltec” 4et50 multi-mode electric transaxle,” *SAE Int. J. Engines*, vol. 4, pp. 1102–1114, 04 2011.
- [41] B. M. Conlon, T. Blohm, M. Harpster, A. Holmes, M. Palardy, S. Tarnowsky, and L. Zhou, “The next generation “voltec” extended range ev propulsion system,” *SAE Int. J. Alt. Power.*, vol. 4, pp. 248–259, 04 2015. [Online]. Available: <https://doi.org/10.4271/2015-01-1152>
- [42] M. Pittel and D. Martin, “effite dedicated hybrid transmission for chrysler pacifica,” in *SAE Technical Paper*. SAE International, 04 2018.
- [43] E. Higuera-Castillo, Z. Kalinic, V. Marinkovic, and F. J. Liébana-Cabanillas, “A mixed analysis of perceptions of electric and hybrid vehicles,” *Energy Policy*, vol. 136, p. 111076, 2020.
- [44] K. Y. Bjerkan, T. E. Nørbech, and M. E. Nordtømme, “Incentives for promoting battery electric vehicle (bev) adoption in norway,” *Transportation Research Part D: Transport and Environment*, vol. 43, pp. 169 – 180, 2016.
- [45] J. S. Krupa, D. M. Rizzo, M. J. Eppstein, D. B. Lanute, D. E. Gaalema, K. Lakkaraju, and C. E. Warrender, “Analysis of a consumer survey on plug-in hybrid electric vehicles,” *Transportation Research Part A: Policy and Practice*, vol. 64, pp. 14 – 31, 2014.



- [46] X. Zhang, X. Bai, and J. Shang, “Is subsidized electric vehicles adoption sustainable: Consumers’ perceptions and motivation toward incentive policies, environmental benefits, and risks,” *Journal of Cleaner Production*, vol. 192, pp. 71 – 79, 2018.
- [47] F. Schmalfuß, K. Mühl, and J. F. Krems, “Direct experience with battery electric vehicles (bevs) matters when evaluating vehicle attributes, attitude and purchase intention,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 46, pp. 47 – 69, 2017.
- [48] R. Halbright and M. Dunn, “Case study: The toyota prius,” *Managerial Marketing*, 2010, [Accessed: 2020-01-02]. [Online]. Available: [https://scalar.usc.edu/works/sustainability-bart-portfolio/media/Prius\\_Marketing\\_Case\\_Study.pdf](https://scalar.usc.edu/works/sustainability-bart-portfolio/media/Prius_Marketing_Case_Study.pdf)
- [49] Plug-in hybrid electric vehicle sales in u.s. by model, 1999 to 2017. [Accessed: 2020-01-05]. [Online]. Available: <https://afdc.energy.gov/data/10576>
- [50] M. M. Sabri, K. Danapalasingam, and M. Rahmat, “A review on hybrid electric vehicles architecture and energy management strategies,” *Renew. Sust. Energ. Rev.*, vol. 53, pp. 1433 – 1442, 2016.
- [51] Y. Gurkaynak, A. Khaligh, and A. Emadi, “State of the art power management algorithms for hybrid electric vehicles,” in *2009 IEEE Veh. Pow. Propuls. Conf.*, Sep. 2009, pp. 388–394.

- [52] S. G. Wirasingha and A. Emadi, “Classification and review of control strategies for plug-in hybrid electric vehicles,” *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 111–122, Jan. 2011.
- [53] B. Moulik and D. Söffker, “Optimal rule-based power management for online, real-time applications in HEVs with multiple sources and objectives: A review,” *Energies*, vol. 8, no. 9, pp. 9049–9063, 2015.
- [54] T. Q. Dinh, J. Marco, H. Niu, D. Greenwood, L. Harper, and D. Corrochano, “A novel method for idle-stop-start control of micro hybrid construction equipment—Part A: Fundamental concepts and design,” *Energies*, vol. 10, no. 7, 2017.
- [55] R. Bojoi, A. Cavagnino, M. Cossale, and A. Tenconi, “Multiphase starter generator for a 48-V mini-hybrid powertrain: Design and testing,” *IEEE Trans. Ind. Appl.*, vol. 52, no. 2, pp. 1750–1758, Mar. 2016.
- [56] J. Seo, S. Kim, and H. Jung, “Rotor-design strategy of IPMSM for 42-V integrated starter generator,” *IEEE Trans. Magn.*, vol. 46, no. 6, pp. 2458–2461, Jun. 2010.
- [57] Y. Zheng, F. Zhao, Y. Luo, and K. Li, “Energy optimization for the mild hybrid electric vehicle with CVT under driving conditions,” in *Proc. of the FISITA 2012 World Automotive Congr.* Springer, 2013, pp. 375–386.
- [58] R. M. Patil, Z. Filipi, and H. K. Fathy, “Comparison of supervisory control strategies for series plug-in hybrid electric vehicle powertrains through dynamic programming,” *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 502–509, Mar. 2014.

- [59] S. D. Cairano, W. Liang, I. V. Kolmanovsky, M. L. Kuang, and A. M. Phillips, “Power smoothing energy management and its application to a series hybrid powertrain,” *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 6, pp. 2091–2103, Nov. 2013.
- [60] Z. Yuan, L. Teng, S. Fengchun, and H. Peng, “Comparative study of dynamic programming and Pontryagin's minimum principle on energy management for a parallel hybrid electric vehicle,” *Energies*, vol. 6, no. 4, pp. 2305–2318, 2013.
- [61] E. Diñçmen and B. A. Güvenç, “A control strategy for parallel hybrid electric vehicles based on extremum seeking,” *Vehicle Syst. Dyn.*, vol. 50, no. 2, pp. 199–227, 2012.
- [62] A. Gao, X. Deng, M. Zhang, and Z. Fu, “Design and validation of real-time optimal control with ECMS to minimize energy consumption for parallel hybrid electric vehicles,” *Math. Probl. Eng.*, vol. 2017, p. 13, 2017.
- [63] A. Rezaei, J. B. Burl, B. Zhou, and M. Rezaei, “A new real-time optimal energy management strategy for parallel hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 830–837, Mar. 2019.
- [64] C. Zhang and A. Vahidi, “Route preview in energy management of plug-in hybrid vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 2, pp. 546–553, Mar. 2012.
- [65] D. Zhao, R. Stobart, G. Dong, and E. Winward, “Real-time energy management for diesel heavy duty hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 3, pp. 829–841, May 2015.

- [66] W. Enang and C. Bannister, “Robust proportional ECMS control of a parallel hybrid electric vehicle,” *Proc. of the Inst. of Mech. Eng., Part D: J. Automob. Eng.*, vol. 231, no. 1, pp. 99–119, 2017.
- [67] F. Khoucha, M. E. H. Benbouzid, and A. Kheloui, “An optimal fuzzy logic power sharing strategy for parallel hybrid electric vehicles,” in *2010 IEEE Veh. Pow. and Propul. Conf.*, Sep. 2010, pp. 1–5.
- [68] V. T. Long and N. V. Nhan, “Bees-algorithm-based optimization of component size and control strategy parameters for parallel hybrid electric vehicles,” *Int. J. Automot. Technol.*, vol. 13, no. 7, pp. 1177–1183, 2012.
- [69] C. Vagg, S. Akehurst, C. J. Brace, and L. Ash, “Stochastic dynamic programming in the real-world control of hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 3, pp. 853–866, May 2016.
- [70] A. A. Malikopoulos, “A multiobjective optimization framework for online stochastic optimal control in hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 2, pp. 440–450, Mar. 2016.
- [71] Y.-H. Cheng and C.-M. Lai, “Control strategy optimization for parallel hybrid electric vehicles using a memetic algorithm,” *Energies*, vol. 10, no. 3, p. 305, 2017.
- [72] N. Denis, M. R. Dubois, J. P. F. Trovão, and A. Desrochers, “Power split strategy optimization of a plug-in parallel hybrid electric vehicle,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 315–326, Jan. 2018.

- [73] F. Yi, Z. Lv, Y. Liu, and H. Liu, “Fuzzy adaptive control strategy with improved PSO algorithm for parallel hybrid electric vehicle,” *Int. J. Simulation–Syst., Sci. & Technol.*, vol. 17, no. 39, 2016.
- [74] J. Liang, J. Zhang, H. Zhang, and C. Yin, “Fuzzy energy management optimization for a parallel hybrid electric vehicle using chaotic non-dominated sorting genetic algorithm,” *Automatika*, vol. 56, no. 2, pp. 149–163, 2015.
- [75] M. Schori, T. J. Boehme, B. Frank, and B. P. Lampe, “Optimal calibration of map-based energy management for plug-in parallel hybrid configurations: A hybrid optimal control approach,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 9, pp. 3897–3907, Sep. 2015.
- [76] A. Taghavipour, N. L. Azad, and J. Mcphee, “An optimal power management strategy for power-split plug-in hybrid electric vehicles,” *Int. J. Vehicle Design*, vol. 604, no. 34, pp. 286–304, 2012.
- [77] D. Shi, S. Wang, P. Pisu, L. Chen, R. Wang, and R. Wang, “Modeling and optimal energy management of a power-split hybrid electric vehicle,” *Sci. China Technol. Sciences*, vol. 60, no. 5, pp. 713–725, 2017.
- [78] S. J. Moura, H. K. Fathy, D. S. Callaway, and J. L. Stein, “A stochastic optimal control approach for power management in plug-in hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 3, pp. 545–555, May 2011.
- [79] Y. L. Murphey, J. Park, L. Kiliaris, M. L. Kuang, M. A. Masrur, A. M. Phillips, and Q. Wang, “Intelligent hybrid vehicle power control—Part II: Online intelligent energy management,” *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 69–79,

Jan. 2013.

- [80] J. Torres, R. Gonzalez, A. Gimenez, and J. Lopez, “Energy management strategy for plug-in hybrid electric vehicles: A comparative study,” *Appl. Energy*, vol. 113, pp. 816 – 824, 2014.
- [81] C. Xiang, F. Ding, W. Wang, and W. He, “Energy management of a dual-mode power-split hybrid electric vehicle based on velocity prediction and nonlinear model predictive control,” *Appl. Energy*, vol. 189, pp. 640 – 653, 2017.
- [82] Y. Cai, M. Ouyang, and F. Yang, “Impact of power-split configurations on fuel consumption and battery degradation in plug-in hybrid electric city buses,” *Appl. Energy*, vol. 188, pp. 257 – 269, 2017.
- [83] D. Zhao, R. Stobart, G. Dong, and E. Winward, “Real-time energy management for diesel heavy duty hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 3, pp. 829–841, May 2015.
- [84] J. Wu, J. Ruan, N. Zhang, and P. D. Walker, “An optimized real-time energy management strategy for the power-split hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, pp. 1–9, 2018.
- [85] L. Xu, M. Ouyang, J. Li, F. Yang, L. Lu, and J. Hua, “Optimal sizing of plug-in fuel cell electric vehicles using models of vehicle performance and system cost,” *Appl. Energy*, vol. 103, pp. 477 – 487, 2013.
- [86] Z. Chen, C. C. Mi, J. Xu, X. Gong, and C. You, “Energy management for a power-split plug-in hybrid electric vehicle based on dynamic programming and

- neural networks,” *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1567–1580, May 2014.
- [87] B. V. Padmarajan, A. McGordon, and P. A. Jennings, “Blended rule-based energy management for PHEV: System structure and strategy,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8757–8762, Oct. 2016.
- [88] X. Zeng and J. Wang, “A parallel hybrid electric vehicle energy management strategy using stochastic model predictive control with road grade preview,” *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 6, pp. 2416–2423, Nov. 2015.
- [89] T. van Keulen, D. van Mullem, B. de Jager, J. T. Kessels, and M. Steinbuch, “Design, implementation, and experimental validation of optimal power-split control for hybrid electric trucks,” *Control Eng. Practice*, vol. 20, no. 5, pp. 547 – 558, 2012.
- [90] S. Stockar, V. Marano, M. Canova, G. Rizzoni, and L. Guzzella, “Energy-optimal control of plug-in hybrid electric vehicles for real-world driving cycles,” *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 2949–2962, Sep. 2011.
- [91] T. Nüesch, A. Cerofolini, G. Mancini, N. Cavina, C. Onder, and L. Guzzella, “Equivalent consumption minimization strategy for the control of real driving NOx emissions of a diesel hybrid electric vehicle,” *Energies*, vol. 7, no. 5, pp. 3148–3178, 2014.
- [92] F. Xu, X. Jiao, Y. Wang, and Y. Jing, “Battery-lifetime-conscious energy management strategy based on SP-SDP for commuter plug-in hybrid electric vehicles,” *IEEE Trans. Electric. and Electron. Eng.*, vol. 13, no. 3, pp. 472–479,

2018.

- [93] L. Tang, G. Rizzoni, and S. Onori, “Energy management strategy for HEVs including battery life optimization,” *IEEE Trans. Transport. Electrific.*, vol. 1, no. 3, pp. 211–222, Oct. 2015.
- [94] D. F. Opila, X. Wang, R. McGee, R. B. Gillespie, J. A. Cook, and J. W. Grizzle, “An energy management controller to optimally trade-off fuel economy and drivability for hybrid vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 6, pp. 1490–1505, Nov. 2012.
- [95] V. Sezer, M. Gokasan, and S. Bogosyan, “A novel ECMS and combined cost map approach for high-efficiency series hybrid electric vehicles,” *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3557–3570, Oct. 2011.
- [96] S. D. Farrall and R. P. Jones, “Energy management in an automotive electric/heat engine hybrid powertrain using fuzzy decision making,” in *Proceedings of 8th IEEE Int. Symp. Intell. Control*, Aug. 1993, pp. 463–468.
- [97] H.-D. Lee and S.-K. Sul, “Fuzzy-logic-based torque control strategy for parallel-type hybrid electric vehicle,” *IEEE Trans. Ind. Electron.*, vol. 45, no. 4, pp. 625–632, Aug. 1998.
- [98] C.-C. Lin, H. Peng, J. W. Grizzle, and J.-M. Kang, “Power management strategy for a parallel hybrid electric truck,” *IEEE Trans. Control Syst. Technol.*, vol. 11, no. 6, pp. 839–849, Nov. 2003.
- [99] A. Piccolo, L. Ippolito, V. zo Galdi, and A. Vaccaro, “Optimisation of energy flow management in hybrid electric vehicles via genetic algorithms,” in



- 2001 IEEE/ASME Int. Conf. Advanced Intell. Mechatronics. Proceedings (Cat. No.01TH8556)*, vol. 1, Jul. 2001, pp. 434–439.
- [100] L. V. Pérez, G. R. Bossio, D. Moitre, and G. O. García, “Optimization of power management in an hybrid electric vehicle using dynamic programming,” *Mathematics and Computers in Simulation*, vol. 73, no. 1, pp. 244 – 254, 2006.
- [101] J. Park, Z. Chen, L. Kiliaris, M. L. Kuang, M. A. Masrur, A. M. Phillips, and Y. L. Murphey, “Intelligent vehicle power control based on machine learning of optimal control parameters and prediction of road type and traffic congestion,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4741–4756, Nov. 2009.
- [102] A. Sciarretta, M. Back, and L. Guzzella, “Optimal control of parallel hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 12, no. 3, pp. 352–363, May 2004.
- [103] D. F. Opila, X. Wang, R. McGee, and J. W. Grizzle, “Real-time implementation and hardware testing of a hybrid vehicle energy management controller based on stochastic dynamic programming,” *J. Dyn. Syst., Mea., Control*, vol. 135, no. 2, p. 021002, 2012.
- [104] T. Nüesch, M. Wang, P. Isenegger, C. H. Onder, R. Steiner, P. Macri-Lassus, and L. Guzzella, “Optimal energy management for a diesel hybrid electric vehicle considering transient PM and quasi-static NOx emissions,” *Control Engineering Practice*, vol. 29, pp. 266 – 276, 2014.

- [105] X. Lin, Y. Wang, P. Bogdan, N. Chang, and M. Pedram, “Reinforcement learning based power management for hybrid electric vehicles,” in *2014 IEEE/ACM Int. Conf. Computer-Aided Design (ICCAD)*, Nov. 2014, pp. 33–38.
- [106] R. E. Bellman, “Dynamic Programming,” *Ann. Oper. Res.*, vol. 11, no. 1, pp. 343–395, 1957.
- [107] L. S. Pontryagin, *Mathematical Theory of Optimal Processes*. Routledge, May 2018.
- [108] H. Son and H. Kim, “Development of near-optimal rule-base control for plug-in hybrid electric vehicles taking into account drivetrain component losses,” *Energies*, vol. 9, no. 6, 2016.
- [109] J. Peng, H. He, and R. Xiong, “Rule-based energy management strategy for a series-parallel plug-in hybrid electric bus optimized by dynamic programming,” *Appl. Energy*, vol. 185, pp. 1633 – 1643, 2017.
- [110] R. Biasini, S. Onori, and G. Rizzoni, “A near-optimal rule-based energy management strategy for medium duty hybrid truck,” *Int. J. Powertrains*, vol. 2, no. 2-3, pp. 232–261, 2013.
- [111] D. Bianchi, L. Rolando, L. Serrao, S. Onori, G. Rizzoni, N. Al-Khayat, T. M. Hsieh, and P. Kang, “A rule-based strategy for a series-parallel hybrid electric vehicle: An approach based on dynamic programming,” in *ASME 2010 Dyn. Syst. Control Conf.*, no. 44175, Sep. 2010, pp. 507–514.
- [112] D. S. Naidu, *Optimal control systems*. CRC press, Aug. 2002.

- [113] C. Hou, M. Ouyang, L. Xu, and H. Wang, “Approximate Pontryagin’s minimum principle applied to the energy management of plug-in hybrid electric vehicles,” *Appl. Energy*, vol. 115, pp. 174 – 189, 2014.
- [114] D. Ambühl, O. Sundström, A. Sciarretta, and L. Guzzella, “Explicit optimal control policy and its practical application for hybrid electric powertrains,” *Control Eng. Pract.*, vol. 18, no. 12, pp. 1429 – 1439, 2010.
- [115] T. Leroy, J. Malaizé, and G. Corde, “Towards real-time optimal energy management of HEV powertrains using stochastic dynamic programming,” in *Proc. IEEE Veh. Power and Propulsion Conf.*, Oct. 2012, pp. 383–388.
- [116] A. Chasse, P. Pognant-Gros, and A. Sciarretta, “Online implementation of an optimal supervisory control for a parallel hybrid powertrain,” *SAE Int. J. Engines*, vol. 2, no. 1, pp. 1630–1638, 2009.
- [117] A. Rezaei, J. B. Burl, and B. Zhou, “Estimation of the ECMS equivalent factor bounds for hybrid electric vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 6, pp. 2198–2205, Nov. 2018.
- [118] M. Debert, G. Colin, Y. Chamailard, L. Guzzella, A. Ketfi-Cherif, and B. Bellicaud, “Predictive energy management for hybrid electric vehicles - prediction horizon and battery capacity sensitivity,” in *Proc. IFAC Symp. Advances Automot. Control*, Munich, Germany, Jul. 2010, p. AAC 2010.
- [119] A. Taghavipour, N. L. Azad, and J. McPhee, “Real-time predictive control strategy for a plug-in hybrid electric powertrain,” *Mechatronics*, vol. 29, pp. 13 – 27, 2015.

- [120] D. F. Opila, X. Wang, R. McGee, R. B. Gillespie, J. A. Cook, and J. W. Grizzle, “Real-world robustness for hybrid vehicle optimal energy management strategies incorporating drivability metrics,” *J. Dyn. Syst., Mea., Control*, vol. 136, no. 6, p. 061011, 2014.
- [121] C. Yang, S. Du, L. Li, S. You, Y. Yang, and Y. Zhao, “Adaptive real-time optimal energy management strategy based on equivalent factors optimization for plug-in hybrid electric vehicle,” *Appl. Energy*, vol. 203, pp. 883 – 896, 2017.
- [122] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd ed. Santa Clara, CA, USA: Springer-Verlag TELOS, 2008.
- [123] S. W. Cheng and R. Janardan, “Algorithms for ray-shooting and intersection searching,” *Journal of Algorithms*, vol. 13, no. 4, pp. 670 – 692, 1992.
- [124] L. Michaels, S. Pagerit, A. Rousseau, P. Sharer, S. Halbach, R. Vijayagopal, M. Kropinski, G. Matthews, M. Kao, O. Matthews, M. Steele, and A. Will, “Model-based systems engineering and control system development via virtual hardware-in-the-loop simulation,” in *SAE Technical Paper*. SAE International, Oct. 2010.
- [125] T. Bourdon, L. Saussol, and B. Varoquié, “Integration of physical AMESim® engine model in hardware in the loop environment, dedicated to engine control unit testing,” in *SAE Technical Paper*. SAE International, Apr. 2007.
- [126] J. Bucher, T. Bradley, H. Lohse-Busch, and E. Rask, “Analyzing the energy consumption variation during chassis dynamometer testing of conventional, hybrid

- electric, and battery electric vehicles,” *SAE Int. J. Alternative Powertrains*, vol. 3, no. 1, pp. 36–43, 2014.
- [127] C. Zhang, A. Vahidi, P. Pisu, X. Li, and K. Tennant, “Role of terrain preview in energy management of hybrid electric vehicles,” vol. 59, no. 3, pp. 1139–1147, March 2010, Accessed: 2017-01-15. [Online]. Available: DOI:10.1109/TVT.2009.2038707
- [128] N. Kim and A. Rousseau, “Sufficient conditions of optimal control based on pontryagin’s minimum principle for use in hybrid electric vehicles,” *Proc. Inst. Mech. Eng. D*, vol. 226, no. 9, pp. 1160–1170, 2012, Accessed: 2018-12-17. [Online]. Available: DOI:10.1177/0954407012438304
- [129] L. Tribioli and S. Onori, “Analysis of energy management strategies in plug-in hybrid electric vehicles: Application to the gm chevrolet volt,” in *Proc. 2013 Am. Control. Conf.*, June 2013, pp. 5966–5971.
- [130] M. Koot, J. T. B. A. Kessels, B. de Jager, W. P. M. H. Heemels, P. P. J. van den Bosch, and M. Steinbuch, “Energy management strategies for vehicular electric power systems,” vol. 54, no. 3, pp. 771–782, May 2005, Accessed: 2019-03-21. [Online]. Available: DOI:10.1109/TVT.2005.847211
- [131] S. M. Fard and A. Khajepour, “An optimal power management system for a regenerative auxiliary power system for delivery refrigerator trucks,” *Appl. Energy*, vol. 169, pp. 748 – 756, 2016, Accessed: 2019-04-17. [Online]. Available: DOI:10.1016/j.apenergy.2016.02.078

- [132] S. Onori, L. Serrao, and G. Rizzoni, “Adaptive equivalent consumption minimization strategy for hybrid electric vehicles,” *Proc. ASME Dyn. Syst. and Control Conf. (DSCC)*, vol. 2010, no. 44175, pp. 499–505, 2010.
- [133] S. Xie, X. Hu, Z. Xin, and J. Brighton, “Pontryagin’s minimum principle based model predictive control of energy management for a plug-in hybrid electric bus,” *Appl. Energy*, vol. 236, pp. 893 – 905, 2019. [Online]. Available: DOI:10.1016/j.apenergy.2018.12.032
- [134] X. Tian, Y. Cai, X. Sun, Z. Zhu, and Y. Xu, “An adaptive ECMS with driving style recognition for energy optimization of parallel hybrid electric buses,” *Energy*, vol. 189, p. 116151, 2019, Accessed: 2019-02-14. [Online]. Available: DOI:10.1016/j.energy.2019.116151
- [135] N. Kim, J. Jeong, and C. Zheng, “Adaptive energy management strategy for plug-in hybrid electric vehicles with pontryagin’s minimum principle based on daily driving patterns,” *Int. J. Pr. Eng. Man.- G. T.*, vol. 6, no. 3, pp. 539–548, 2019. [Online]. Available: DOI:10.1007/s40684-019-00046-z
- [136] Q. Jiang, F. Ossart, and C. Marchand, “Comparative study of real-time hev energy management strategies,” vol. 66, no. 12, pp. 10 875–10 888, Dec 2017, Accessed: 2019-12-30. [Online]. Available: DOI:10.1109/TVT.2017.2727069
- [137] X. Jiao, Y. Li, F. Xu, and Y. Jing, “Real-time energy management based on ECMS with stochastic optimized adaptive equivalence factor for HEVs,” *Cogent Engineering*, vol. 5, no. 1, p. 1540027, 2018, Accessed: 2019-02-17. [Online]. Available: DOI:10.1080/23311916.2018.1540027

- [138] S. Xie, X. Hu, S. Qi, and K. Lang, “An artificial neural network-enhanced energy management strategy for plug-in hybrid electric vehicles,” *Energy*, vol. 163, pp. 837 – 848, 2018. [Online]. Available: DOI:10.1016/j.energy.2018.08.139
- [139] Y. Du, Y. Zhao, Q. Wang, Y. Zhang, and H. Xia, “Trip-oriented stochastic optimal energy management strategy for plug-in hybrid electric bus,” *Energy*, vol. 115, pp. 1259 – 1271, 2016.
- [140] A. Stotsky, “Adaptive estimation of the engine friction torque,” *Eur. J. Control*, vol. 13, no. 6, pp. 618 – 624, 2007.
- [141] A. Biswas, R. Gu, P. Kollmeyer, R. Ahmed, and A. Emadi, “Simultaneous state and parameter estimation of li-ion battery with one state hysteresis model using augmented unscented kalman filter,” in *Proc. 2018 IEEE Transport. Electric, Conf. Expo (ITEC)*, June 2018, pp. 1065–1070.
- [142] W. Zhuang, X. Zhang, D. Li, L. Wang, and G. Yin, “Mode shift map design and integrated energy management control of a multi-mode hybrid electric vehicle,” *Appl. Energy*, vol. 204, pp. 476 – 488, 2017.
- [143] T. C. J. Romijn, M. C. F. Donkers, J. T. B. A. Kessels, and S. Weiland, “A distributed optimization approach for complete vehicle energy management,” vol. 27, no. 3, pp. 964–980, May 2019, Accessed: 2020-04-21. [Online]. Available: DOI:10.1109/TCST.2018.2789464
- [144] T. Romijn, T. Pham, and S. Wilkins, “Modular ECMS framework for hybrid vehicles,” *IFAC-Papers OnLine*, vol. 52, no. 5, pp. 128 – 133, 2019.

- [145] T. Hester, M. Quinlan, and P. Stone, “Rtmba: A real-time model-based reinforcement learning architecture for robot control,” in *2012 IEEE Int. Conf. Robotics and Automation*, May 2012, pp. 85–90.
- [146] Y. Nakamura, T. Mori, M. aki Sato, and S. Ishii, “Reinforcement learning for a biped robot based on a cpg-actor-critic method,” *Neural Networks*, vol. 20, no. 6, pp. 723 – 735, 2007.
- [147] Q. Wang and C. Phillips, “Cooperative collision avoidance for multi-vehicle systems using reinforcement learning,” in *2013 18th Int. Conf. Methods Models in Automation Robotics (MMAR)*, Aug. 2013, pp. 98–102.
- [148] W. Usaha and J. A. Barria, “Reinforcement learning for resource allocation in leo satellite networks,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 37, no. 3, pp. 515–527, 2007.
- [149] B. Schölkopf, J. Platt, and T. Hofmann, *Natural Actor-Critic for Road Traffic Optimisation*, 2007, pp. 1169–1176.
- [150] A. Vogel, D. Ramachandran, R. Gupta, and A. Raux, “Improving hybrid vehicle fuel efficiency using inverse reinforcement learning,” in *Proc. of the 26th AAAI Conf. on Artif. Intell.*, 2012, pp. 384–390.
- [151] A. Biswas, P. G. Anselma, and A. Emadi, “Real-time optimal energy management of electrified powertrains with reinforcement learning,” in *Proc. 2019 IEEE Transp. Electrification Conf. (ITEC)*, June 2019, pp. 1–6.



- [152] Y. Fang, C. Song, B. Xia, and Q. Song, “An energy management strategy for hybrid electric bus based on reinforcement learning,” in *The 27th Chinese Control and Decision Conf. (2015 CCDC)*, May 2015, pp. 4973–4977.
- [153] X. Lin, P. Bogdan, N. Chang, and M. Pedram, “Machine learning-based energy management in a hybrid electric vehicle to minimize total operating cost,” in *2015 IEEE/ACM Int. Conf. Computer-Aided Design (ICCAD)*, Nov. 2015, pp. 627–634.
- [154] X. Qi, G. Wu, K. Boriboonsomsin, M. J. Barth, and J. Gonder, “Data-driven reinforcement learning-based real-time energy management system for plug-in hybrid electric vehicles,” *Transp. Res. Rec.*, vol. 2572, no. 1, pp. 1–8, 2016.
- [155] Y. Li, H. He, J. Peng, and H. Zhang, “Power management for a plug-in hybrid electric vehicle based on reinforcement learning with continuous state and action spaces,” *Energy Procedia*, vol. 142, pp. 2270 – 2275, 2017.
- [156] Y. Hu, W. Li, K. Xu, T. Zahid, F. Qin, and C. Li, “Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning,” *Appl. Sci.*, vol. 8, no. 2, 2018.
- [157] J. Wu, H. He, J. Peng, Y. Li, and Z. Li, “Continuous reinforcement learning of energy management with deep q network for a power split hybrid electric bus,” *Appl. Energy*, vol. 222, pp. 799 – 811, 2018.
- [158] R. Xiong, J. Cao, and Q. Yu, “Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle,” *Appl. Energy*, vol. 211, pp. 538 – 548, 2018.

- [159] T. Liu, B. Wang, and C. Yang, “Online Markov chain-based energy management for a hybrid tracked vehicle with speedy Q-learning,” *Energy*, vol. 160, pp. 544 – 555, 2018.
- [160] X. Qi, Y. Luo, G. Wu, K. Boriboonsomsin, and M. Barth, “Deep reinforcement learning enabled self-learning control for energy efficient driving,” *Transp. Res. Part C Emerg. Technol.*, vol. 99, pp. 67 – 81, 2019.
- [161] B. Xu, D. Rathod, D. Zhang, A. Yebi, X. Zhang, X. Li, and Z. Filipi, “Parametric study on reinforcement learning optimized energy management strategy for a hybrid electric vehicle,” *Applied Energy*, vol. 259, p. 114200, 2020.
- [162] X. Qi, G. Wu, K. Boriboonsomsin, and M. J. Barth, “A novel blended real-time energy management strategy for plug-in hybrid electric vehicle commute trips,” in *2015 IEEE 18th Int. Conf. Intel. Transp. Syst.*, Sep. 2015, pp. 1002–1007.
- [163] Z. Kong, Y. Zou, and T. Liu, “Implementation of real-time energy management strategy based on reinforcement learning for hybrid electric vehicles and simulation validation,” *PLOS ONE*, vol. 12, no. 7, pp. 1–16, 07 2017.
- [164] J. N. Tsitsiklis and B. van Roy, “Feature-based methods for large scale dynamic programming,” *Machine Learning*, vol. 22, no. 1, pp. 59–94, 1996.
- [165] J. N. Tsitsiklis and B. Van Roy, “An analysis of temporal-difference learning with function approximation,” vol. 42, no. 5, pp. 674–690, May 1997.
- [166] P. C. Kainen, V. Kůrková, and A. Vogt, “Approximation by neural networks is not continuous,” *Neurocomputing*, vol. 29, no. 1, pp. 47 – 56, 1999.

- [167] V. Kurkova and M. Sanguineti, “Comparison of worst case errors in linear and neural network approximation,” *IEEE Trans. Inf. Theory*, vol. 48, no. 1, pp. 264–275, Jan 2002.
- [168] I. Grondman, L. Busoniu, G. A. D. Lopes, and R. Babuska, “A survey of actor-critic reinforcement learning: Standard and natural policy gradients,” *IEEE Trans on Syst., Man, and Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1291–1307, 2012.
- [169] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. A. Riedmiller, “Playing Atari with deep reinforcement learning,” *CoRR*, vol. abs/1312.5602, 2013.
- [170] M. Riedmiller *et al.*, “Neural fitted Q iteration – First experiences with a data efficient neural reinforcement learning method,” in *Machine Learning: ECML 2005*, Berlin, Heidelberg, 2005, pp. 317–328.
- [171] L.-J. Lin, “Self-improving reactive agents based on reinforcement learning, planning and teaching,” *Machine Learning*, vol. 8, no. 3-4, pp. 293–321, May 1992.
- [172] L. J. Lin, “Reinforcement learning for robots using neural networks,” Ph.D. dissertation, Carnegie Mellon University, USA, 1992.
- [173] V. Mnih *et al.*, “Asynchronous methods for deep reinforcement learning,” in *Proc. 33rd Int. Conf. on Machine Learning*, vol. 48, New York, USA, 20–22 Jun 2016, pp. 1928–1937.

- [174] R. J. Williams, “Simple statistical gradient-following algorithms for connectionist reinforcement learning,” *Machine Learning*, vol. 8, no. 3, pp. 229–256, May 1992.
- [175] X. Qi, Y. Luo, G. Wu, K. Boriboonsomsin, and M. Barth, “Deep reinforcement learning enabled self-learning control for energy efficient driving,” *Transp. Res. Part C Emerg. Technol.*, vol. 99, pp. 67 – 81, 2019.
- [176] X. Zeng, N. Yang, J. Wang, D. Song, N. Zhang, M. Shang, and J. Liu, “Predictive-model-based dynamic coordination control strategy for power-split hybrid electric bus,” *Mech. Syst. Signal Process.*, vol. 60-61, pp. 785 – 798, 2015.
- [177] Z. Zhao and X. Tang, “Control of e-cvt to ev mode transition on braking deceleration condition for a compound power-split hybrid electrical vehicle,” *IEEE Access*, vol. 7, pp. 111 954–111 967, 2019.
- [178] R. J. Williams and J. Peng, “Function optimization using connectionist reinforcement learning algorithms,” *Connection Science*, vol. 3, no. 3, pp. 241–268, 1991.
- [179] T. Zahavy, Z. Xu, V. Veeriah, M. Hessel, J. Oh, H. van Hasselt, D. Silver, and S. Singh, “A self-tuning actor-critic algorithm,” *arXiv preprint arXiv:2002.12928*, 2020.
- [180] W. Zhuang, X. Zhang, Y. Ding, L. Wang, and X. Hu, “Comparison of multi-mode hybrid powertrains with multiple planetary gears,” *Appl. Energy*, vol. 178, pp. 624 – 632, 2016.

- [181] P. G. Anselma, Y. Huo, J. Roeleveld, G. Belingardi, and A. Emadi, “Slope-weighted energy-based rapid control analysis for hybrid electric vehicles,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4458–4466, May 2019.
- [182] C. Wang, Z. Zhao, T. Zhang, and M. Li, “Mode transition coordinated control for a compound power-split hybrid car,” *Mech. Syst. Signal Process.*, vol. 87, pp. 192 – 205, 2017.
- [183] C. Yang, X. Jiao, L. Li, Y. Zhang, and Z. Chen, “A robust h-infinity control-based hierarchical mode transition control system for plug-in hybrid electric vehicle,” *Mech. Syst. Signal Process.*, vol. 99, pp. 326 – 344, 2018.
- [184] X. Wang, L. Li, and C. Yang, “Hierarchical control of dry clutch for engine-start process in a parallel hybrid electric vehicle,” *IEEE Trans. Transp. Electrification*, vol. 2, no. 2, pp. 231–243, June 2016.
- [185] C. Yang, X. Jiao, L. Li, Y. Zhang, L. Zhang, and J. Song, “Robust coordinated control for hybrid electric bus with single-shaft parallel hybrid powertrain,” *IET Control Theory Appl.*, vol. 9, no. 2, pp. 270–282, 2015.
- [186] H. Zhang, Y. Zhang, and C. Yin, “Hardware-in-the-loop simulation of robust mode transition control for a series-parallel hybrid electric vehicle,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1059–1069, March 2016.
- [187] Z. Lei, D. Sun, Y. Liu, D. Qin, Y. Zhang, Y. Yang, and L. Chen, “Analysis and coordinated control of mode transition and shifting for a full hybrid electric

- vehicle based on dual clutch transmissions,” *Mech. Mach. Theory*, vol. 114, pp. 125 – 140, 2017.
- [188] X. Tang, D. Zhang, T. Liu, A. Khajepour, H. Yu, and H. Wang, “Research on the energy control of a dual-motor hybrid vehicle during engine start-stop process,” *Energy*, vol. 166, pp. 1181 – 1193, 2019.
- [189] K. Ahn and S. W. Cha, “Developing mode shift strategies for a two-mode hybrid powertrain with fixed gears,” *SAE Int. J. Passeng. Cars – Mech. Syst.*, vol. 1, pp. 285–292, 04 2008. [Online]. Available: <https://doi.org/10.4271/2008-01-0307>
- [190] D. Karbowski, J. Kwon, N. Kim, and A. Rousseau, “Instantaneously optimized controller for a multimode hybrid electric vehicle,” in *SAE Technical Paper*. SAE International, 04 2010.
- [191] W. Zhuang, X. Zhang, G. Yin, H. Peng, and L. Wang, “Mode shift schedule and control strategy design of multimode hybrid powertrain,” *IEEE Tran. Control Syst. Technol.*, vol. 28, no. 3, pp. 804–815, May 2020.
- [192] W. Geng, D. Lou, C. Wang, and T. Zhang, “A cascaded energy management optimization method of multimode power-split hybrid electric vehicles,” *Energy*, vol. 199, p. 117224, 2020.
- [193] P. G. Anselma, Y. Huo, J. Roeleveld, G. Belingardi, and A. Emadi, “Integration of on-line control in optimal design of multimode power-split hybrid electric vehicle powertrains,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3436–3445, April 2019.

- [194] G. Buccoliero, P. G. Anselma, S. A. Bonab, G. Belingardi, and A. Emadi, “A new energy management strategy for multimode power-split hybrid electric vehicles,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 172–181, Jan 2020.
- [195] X. Zhang, S. Eben Li, H. Peng, and J. Sun, “Efficient exhaustive search of power-split hybrid powertrains with multiple planetary gears and clutches,” *J. Dyn. Syst. Meas. Control*, vol. 137, no. 12, 09 2015, 121006.
- [196] P. G. Anselma, A. Biswas, G. Belingardi, and A. Emadi, “Rapid assessment of the fuel economy capability of parallel and series-parallel hybrid electric vehicles,” *Appl. Energy*, vol. 275, p. 115319, 2020.
- [197] F. U. Syed, M. L. Kuang, J. Czubay, and H. Ying, “Derivation and experimental validation of a power-split hybrid electric vehicle model,” *IEEE Transactions on Vehicular Technology*, vol. 55, no. 6, pp. 1731–1747, Nov 2006.
- [198] J. Wang, Y. Cai, L. Chen, D. Shi, R. Wang, and Z. Zhu, “Review on multi-power sources dynamic coordinated control of hybrid electric vehicle during driving mode transition process,” *Int. J. Energy Res.*, vol. 44, no. 8, pp. 6128–6148, 2020.
- [199] L. Chen, G. Xi, and J. Sun, “Torque coordination control during mode transition for a series-parallel hybrid electric vehicle,” *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 2936–2949, Sep. 2012.
- [200] K. D. Mishra and K. Srinivasan, “Robust control and estimation of clutch-to-clutch shifts,” *Control Eng. Pract.*, vol. 65, pp. 100 – 114, 2017.

- 
- [201] P. G. Anselma, A. Biswas, J. Roeleveld, G. Belingardi, and A. Emadi, “Multi-fidelity near-optimal on-line control of a parallel hybrid electric vehicle powertrain,” in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, June 2019, pp. 1–6.
- [202] P. Dekraker, M. Stuhldreher, and Y. Kim, “Characterizing factors influencing si engine transient fuel consumption for vehicle simulation in alpha,” *SAE Int. J. Engines*, vol. 10, pp. 529–540, 03 2017.