



Energy and Service Life Management Strategy for a Two-Drive Multi-Speed Electric Vehicle

Am Fachbereich Maschinenbau

an der Technischen Universität Darmstadt

zur

Erlangung des Grades eines Doktor-Ingenieurs (Dr.-Ing.)

genehmigte

DISSERTATION

vorgelegt von

YIKAI TAO, M.SC.

aus Shanghai, China

Berichterstatter: Prof. Dr.-Ing. Stephan Rinderknecht

Mitberichterstatter: Prof. Dr.-Ing. Christian Beidl

Tag der Einreichung: 06.03.2024

Tag der mündlichen Prüfung: 17.04.2024

Darmstadt 2024

D17

Yikai, Tao : Energy and Service Life Management Strategy for a Two-Drive Multi-Speed Electric Vehicle

Darmstadt, Technische Universität Darmstadt,

Jahr der Veröffentlichung der Dissertation auf TUprints: 2024

Tag der mündlichen Prüfung: 17.04.2024

Veröffentlicht unter CC BY 4.0 International

<https://creativecommons.org/licenses/>

Abstract

Regulations of zero emission passenger cars appear on the horizon, and battery electric vehicles (BEV) are the main solution from the current market. It has been a focus of both academia and industry to extend their range. One of the main approaches is to reduce their energy consumption. Recent studies have shown that the two-drive topology and the multi-speed topology help to do so. It is natural to combine both concepts and to design a two-drive multi-speed topology for BEVs. Due to its more than one degree of freedom, an online energy management strategy (EMS) controlling torque set points of both electric motors and target gear positions is necessary to exploit its potential for reducing total energy consumption in real-world applications. There are numerous studies on EMSs for BEVs and hybrid electric vehicles. The overwhelming majority of them shared the same assumption: shift processes are neglectable. Based on the shift duration statistics, the shift processes of the most common transmissions in today's market are too long to be ignored for an EMS with an operation frequency of at least 1 Hz. How to develop an EMS that considers shift processes? Suppose that an EMS is developed. It controls the powertrain in favour of low energy consumption, and the parts and the components are loaded accordingly. Some parts might fatigue and fail much faster than others, not because of poor construction dimensioning, but because of excessive use. What can an EMS do to prevent such an extreme scenario? Furthermore, is there a general way to design EMSs for multi-drive BEVs?

This thesis is initiated by developing an online EMS for a two-drive multi-speed BEV called "Speed4E", and tends to address the questions raised earlier. A predictive EMS in a Model Predictive Control framework is developed. A hybrid system considering the shift processes is proposed. Based on it and the Hybrid Minimum Principle, a solver and its algorithms are developed. The Principle is chosen for its accuracy and low time complexity, the two most important attributes of an online EMS. Minimizing the instantaneous Hamiltonian in the Principle is mathematically analysed. Several Lemmas that reduce the time complexity considerably are produced. Compared to an EMS that minimizes instantaneous energy consumption and ignores shift processes, the predictive EMS reduces the energy consumption in the Worldwide Harmonized Light Vehicles Test Cycle (WLTC) by 0.26 % and the shift count by 63.41 %. The hybrid system, the predictive EMS and the mathematical analysis are, as far as the author knows, first of their kinds. A novel multi-criteria operation strategy (MCOS) considering powertrain service life is proposed. Thanks to the hybrid system, the influence of the shift processes on fatigue is included. The MCOS extends the powertrain service life by several times but sacrifices the energy consumption. A general multi-drive (at least two) multi-speed electric powertrain is proposed. Its hybrid system is formulated. The Principle is applied to produce the optimality condition. It is showcased, how to modify certain sets and sample space in the formulation to have the general model and problem represent certain electric powertrains. A unified framework to design EMS for the general multi-drive electric powertrain is proposed, where the algorithms developed for the predictive EMS can be applied.



Erklärung zur Dissertation

Hiermit versichere ich, die vorliegende Dissertation ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 05.03.2024

Contents

| | |
|---|-----------|
| Abstract..... | I |
| Erklärung zur Dissertation..... | II |
| Contents..... | III |
| List of Figures..... | VI |
| List of Tables..... | XII |
| Nomenclature..... | XIII |
| <i>Acronyms and Abbreviations</i> | XIII |
| <i>Roman Symbols</i> | XIV |
| <i>Greek Symbols</i> | XVII |
| <i>Superscripts</i> | XVIII |
| <i>Subscripts</i> | XVIII |
| 1 Introduction..... | 1 |
| 1.1 Motivation..... | 2 |
| 1.2 Goals and Contributions..... | 5 |
| 1.3 Structure of the Content..... | 6 |
| 1.4 Acronyms and Abbreviations..... | 6 |
| 2 Background Knowledge..... | 8 |
| 2.1 Hybrid Systems and Regularity Conditions..... | 8 |
| 2.2 Formulation of Hybrid Optimal Control Problems..... | 14 |
| 2.3 Discretization and Integration Schemes..... | 18 |
| 2.4 State of the Art Solution Methods..... | 20 |
| 2.5 Summary..... | 32 |
| 3 State of the Art..... | 33 |
| 3.1 Topology and Optimal Control Problem..... | 33 |
| 3.2 Energy Management Strategies..... | 36 |
| 3.3 Multi-Criteria Operation Strategies..... | 41 |
| 3.4 Summary and Other Aspects..... | 42 |
| 4 Hybrid System and Problem Formulation..... | 44 |
| 4.1 Speed4E Powertrain..... | 44 |
| 4.2 Longitudinal Dynamics..... | 45 |
| 4.3 Shift Processes..... | 47 |
| 4.4 Hybrid System Formulation..... | 49 |

| | | |
|-------------------|--|------------|
| 4.5 | Hybrid Optimal Control Problem of Minimal Energy Consumption | 56 |
| 4.6 | Summary | 57 |
| 5 | Development of Energy Management Strategy..... | 59 |
| 5.1 | Vehicle Speed Prediction..... | 59 |
| 5.2 | Hybrid Minimum Principle Solution Method | 61 |
| 5.3 | Minimization of the Cost Function | 69 |
| 5.4 | Time Complexity of the Predictive EMS..... | 75 |
| 5.5 | Summary..... | 78 |
| 6 | Simulation Results of the Predictive Energy Management Strategy | 80 |
| 6.1 | Reference Energy Management Strategies | 80 |
| 6.2 | Solution Process of the Predictive EMS..... | 81 |
| 6.3 | Driving Cycle Simulation..... | 91 |
| 6.4 | Summary..... | 101 |
| 7 | General Multi-Drive Multi-Speed Electric Vehicle..... | 104 |
| 7.1 | Hybrid System Formulation | 104 |
| 7.2 | Hybrid Optimal Control of the General Multi-Drive Multi-Speed Electric Vehicle . | 110 |
| 7.3 | Summary..... | 114 |
| 8 | Multi-Criteria Operation Strategy Considering Service Life | 116 |
| 8.1 | Service Life Estimation | 116 |
| 8.2 | Development of the Multi-Criteria Operation Strategy..... | 134 |
| 8.3 | Effect of Energy Management Strategies on Service Life | 139 |
| 8.4 | Driving Cycle Simulation..... | 146 |
| 8.5 | Summary..... | 158 |
| 9 | Conclusion and Outlook..... | 161 |
| Appendix A | The Parameters of Speed4E Powertrain and the vehicle | 167 |
| Appendix B | Speed4E Powertrain Power Losses | 170 |
| Appendix C | Dynamics of a M-Drive Multi-Speed Powertrain | 172 |
| Appendix D | Cost Function Minimization..... | 181 |
| Appendix E | The Bearings and Gears in Speed4E Powertrain | 190 |
| Appendix F | Multi-Criteria HOCF with Option 3..... | 195 |
| Appendix G | Fatigue Coefficient | 197 |
| Appendix H | Accumulated Fatigue and Service Life | 198 |



References 202

Standards..... 218

List of Figures

Figure 1.1: EV sales volumes [4]. (a) EV worldwide sales from 2016 to 2023. (b) BEV and PHEV sales in EU from 2018 to 2022. 1

Figure 1.2: EMS as high-level control in a typical BEV powertrain control scheme..... 3

Figure 2.1: A bouncing ball..... 11

Figure 2.2: Hybrid automaton of a bouncing ball..... 12

Figure 2.3: RK4 integration scheme performed on a 1-D system, whose dynamic is $\dot{x} = 0.4x + 0.1$. Analytical solution is $xt = e0.4t + C - 0.04$ with $C = 0$, $tk = 1$ and $h = 0.1$ 20

Figure 2.4: An example of discrete Bellman equation on a discretization grid $Gx \times Q \times Gt$. A transition to discrete state $q2$ or $q3$ is not admissible, which is defined by the admissible discrete control set $Bq1$ 22

Figure 2.5: An example of a solution with HMP to a HOCP with a 1-D continuous state. (a) shows continuous state trajectory. (b) shows costate trajectory..... 25

Figure 2.6: Schematic of shooting method to solve a TPBVP transformed from a HOCP of 1-D continuous state by HMP. (a) continuous state trajectories in different iterations. (b) costate trajectories in different iterations. 27

Figure 2.7: Schematic of direct shooting method to a HOCP with 1-D continuous state to MIP. (a) 1-D continuous state and discrete state. (b) Piecewise control out of discretized continuous control grid Gu 29

Figure 3.1. Typical powertrain topologies of BEVs. (a): BEV with in-wheel motors. (b) Single-drive BEV. (c): Centralized two-drive BEV. (d): Separated two-drive BEV T: Transmission. ST: Sub-transmission. FD: Final Drive. 34

Figure 3.2: Schematic of a planetary gearbox..... 35

Figure 3.3: Schematic of MPC 39

Figure 3.4. Typical components of a predictive EMS for continuous OCPs. 40

Figure 4.1: Advantages of high-speed e-drives w.t.r. weight of active components [139]. 44

Figure 4.2: Topology of Speed4E powertrain [141] 45

Figure 4.3: Shift process from 1st to 2nd gear position in the EMS..... 47

Figure 4.4: Shift process in the EMS. (a): from 1st to neutral gear position. (b): from neutral to 1st gear position. 48

Figure 4.5: Hybrid automaton of Speed4E powertrain..... 50

Figure 5.1: Basic structure of the predictive EMS embedded with the HMP solution method. 59

Figure 5.2: WLTC speed profile..... 60

Figure 5.3: Flowchart of shootHMP 65

Figure 5.4: Flowchart of solveHMP..... 67

Figure 5.5: Flowchart of genInitialCOSTATE 67

Figure 5.6: An example of the initialization for a predicted speed profile with constant acceleration

and the system is in $q2$. (a): predicted X and T_{total} . (b): approximated $U1_{alone}$ and $U2_{alone}$. Compare the cost function over the prediction horizon. (c): form $U1$ and $U2$. (d) estimate $\lambda0, U1, \lambda0, U2$ by backwards integration..... 68

Figure 5.7: The constraints of the minimization problem with following parameters: $v = 60$ km/h, $T_{total} = 1000$ Nm, $q = q1$. (a) problem (5.33); (b) equivalent problem. 69

Figure 5.8: Torque range divided into several zones..... 71

Figure 5.9: Illustration for Proof of Lemma 2..... 73

Figure 5.10: Structure of the predictive EMS embedded with the HMP solution method, supplemented by Lemma1, 2 and 3. TPBVP solved by the Broyden method with the Bang-Bang controls inspired initialization..... 79

Figure 6.1: Target gear position look-up table..... 81

Figure 6.2: Predicted speed of individual solutions. (a) $t = 512$ s in the WLTC. (b) $t = 376$ s in the WLTC. (c) high speed and moderate acceleration. 82

Figure 6.3: The instantaneous Hamiltonian in the 1st second: (a) angular velocities, (b) torques, (c) costates, (d) cost function and (e) Hamiltonian. 83

Figure 6.4: Comparison of the solution by minimizing the instantaneous cost, (a) and (b), to the one by minimizing the instantaneous Hamiltonian, (c) and (d). (a) and (c): states and costates. (b) and (d): continuous and discrete controls. (e) Deviation of cumulated cost. 84

Figure 6.5: Hamiltonian given different costates in 1st second. (a) costates. (b) Hamiltonian. 85

Figure 6.6: Comparison of the solution by minimizing instantaneous cost, (a) and (b), to the one by minimizing instantaneous Hamiltonian, (c) and (d). (a) and (c): states and costates. (b) and (d): continuous and discrete controls. (e) Deviation of cumulated cost. 86

Figure 6.7: Compare the solution by minimizing instantaneous cost, (a) and (b), to the one by minimizing instantaneous Hamiltonian, (d) and (e). (a) and (d): states and costates. (b) and (e): continuous and discrete controls. (c) deviation of vehicle speed. (f) deviation of cumulated cost..... 87

Figure 6.8: Convergence behaviour for the speed profile in Figure 6.2(a). (a) The error of the transversality condition at each iteration. (b) The initial costates vs. the Euclidean norm of errors at each iteration 88

Figure 6.9: $|\Phi1|$ and $|\Phi2|$ of the solution process for the speed profile in Figure 6.2(b). 89

Figure 6.10: $|\Phi1|$ and $|\Phi2|$ of the solution process for the speed profile in Figure 6.2(c). 89

Figure 6.11: Structure of the predictive EMS embedded with HMP solution method, whose TPBVP solved by SQP. 90

Figure 6.12: Comparison of the iterations of error..... 90

Figure 6.13: Iteration numbers for solutions to converge. (a): solveHMP with the Broyden method. (b): solveHMP with the SQP. 91

Figure 6.14: Discrete state sequences of the EMS naive optimal and the predictive EMS..... 92

Figure 6.15: Predicted speed vs. real speed at (a) 663 s and (b) 797 s. 92

| | |
|---|-----|
| Figure 6.16: Gear distribution of (a) the EMS naive optimal and (b) the predictive EMS. | 92 |
| Figure 6.17: Control and energy consumption between 1560 s and 1580 s. (a) Speed and acceleration. (b) Energy consumption and its deviation. (c) Torques and gear positions of the EMS naive optimal. (d) Torque and gear positions of the predictive EMS. | 93 |
| Figure 6.18: Operating duration of both EMs. (a) and (b): the EM1 and the EM2 of the EMS naive optimal. (c) and (d): the EM1 and the EM2 of the predictive EMS. | 94 |
| Figure 6.19: Control and energy consumption between 255 s and 295 s. (a) Speed and acceleration. (b) Energy consumption and its deviation. (c) Torques and gear positions of the EMS naive optimal. (d) Torque and gear positions of the predictive EMS. | 95 |
| Figure 6.20: Discrete state sequences of the EMS global optimal and the predictive EMS. | 95 |
| Figure 6.21: Gear distribution of the EMS global optimal. | 96 |
| Figure 6.22: Control and energy consumption between 1150 s and 1170 s. (a) Speed and acceleration. (b) Energy consumption and its deviation. (c) Torques and gear positions of the EMS naive optimal. (d) Torque and gear positions of the predictive EMS. | 97 |
| Figure 6.23: Energy consumed and recuperated during shifts between 1150 s and 1170 s. | 97 |
| Figure 6.24: Predicted speed vs. real speed between 1153 s (0 s in the figure) and 1163 s (10 s in the figure)..... | 97 |
| Figure 6.25: Structure of the predictive EMS with zero-error prediction. | 98 |
| Figure 6.26: Discrete state sequences of the EMS global optimal and the predictive EMS with zero- error prediction | 99 |
| Figure 6.27: Energy consumptions and shift counts of the predictive EMS with different prediction horizon. | 99 |
| Figure 6.28: WLTC Simulation elapsed time of the predictive EMS with different prediction horizon. | 100 |
| Figure 6.29: Average iteration numbers of the predictive EMS with different prediction horizon... | 100 |
| Figure 6.30: Iteration numbers for solutions in the predictive EMS to converge. Np = 30 | 101 |
| Figure 6.31: Energy consumptions and shift counts of all mentioned driving cycle simulation. | 102 |
| Figure 6.32: WLTC Simulation elapsed time of the predictive EMS. | 102 |
| Figure 6.33: Discrete state sequences of the predictive EMS and the EMS naive optimal. | 103 |
| Figure 6.34: Gear distribution of (a) the EMS naive optimal and (b) the predictive EMS. | 103 |
| Figure 7.1: Hybrid automaton of a M-drive multi-speed powertrain with the first assumption. | 105 |
| Figure 7.2: Hybrid automaton of a multi-drive multi-speed powertrain with the second assumption. | 107 |
| Figure 7.3: Hybrid automaton of a multi-drive multi-speed powertrain with the third assumption. | 108 |
| Figure 8.1: The bathtub curve | 117 |
| Figure 8.2: Proportion of time that different gears are engaged, when the vehicle drives with an example driving cycle. | 118 |

| | |
|--|-----|
| Figure 8.3: Typical damage examples. (left) Macropitting. (right) Tooth root breakage— tooth failure [161] | 120 |
| Figure 8.4: Spalling on the inner ring of a rolling bearing [167]..... | 121 |
| Figure 8.5: Schematic of the bending stress on the tooth root. | 122 |
| Figure 8.6: Schematic of the Hertzian stress on the tooth flanks | 123 |
| Figure 8.7: Risk volume of rolling bearing fatigue suggested by [113]. Figure taken from [114]. | 124 |
| Figure 8.8: S-N curves of different Bx -service life of an example part for a certain failure. Figure modified from the one from [32, p. 254]..... | 124 |
| Figure 8.9: Bearing and gaset configuration of Speed4E powertrain. | 129 |
| Figure 8.10: Modified S-N Curves of (a) B1.1.1, (b) G1.S1.SG pitting and (c) G1.S1.SG tooth root breakage. | 130 |
| Figure 8.11: Schematic of meshing force | 130 |
| Figure 8.12: An example of forces on bearings..... | 131 |
| Figure 8.13: Accumulated fatigue after a WLTC and corresponding service lives of the bearings and gears in Speed4E powertrain. (a): Bearing fatigue. (b): Tooth root breakage. (c) Tooth flank pitting. | 132 |
| Figure 8.14: Load and stress spectra under the WLTC of (a) B1.1.2, (b) G1.S1.SG breakage, (c) G1.S1.SG pos pitting, (d) G1.S1.SG neg pitting, (e) B2.1.3, (f) G2.S1.P breakage, (g) G2.S1.P pos pitting, and (h) G2.S1.P neg pitting. | 133 |
| Figure 8.15: Fatigue rate over angular velocity of (a): gears, (b): bearings and (c): bearings zoom in. The total output torque is 1000 Nm. The ST2 in the 1 st gear position..... | 140 |
| Figure 8.16: The average and the larger values of dv of G1.S1.SG pos and G2.S1.P pos. The total output torque is 1000 Nm. The ST2 in the 1 st gear position..... | 141 |
| Figure 8.17: The average and the larger values of dv of B1.1.2 and B2.1.3. The total output torque is 1000 Nm. The ST2 in the 1 st gear position. | 142 |
| Figure 8.18: Fatigue rate over angular velocity of (a): gears, (b): bearings and (c): bearings zoom in. The total output torque is 1000 Nm. The ST2 in the 2 nd gear position..... | 143 |
| Figure 8.19: The average and the larger values of dv . (a): G1.S1.SG pos and G2.S1.P pos. (b): B1.1.2 and B2.1.3 | 143 |
| Figure 8.20: Two discrete state sequences..... | 144 |
| Figure 8.21: The continuous controls, the gear position and the continuous states between 12 s and 18 s. (a) and (b): the strategy 1. (c) and (d): the strategy 2. | 144 |
| Figure 8.22: The Hertzian stresses on the gears and the resulting fatigue between 12 s and 18 s. (a) and (b): the strategy 1. (c) and (d): the strategy 2. | 145 |
| Figure 8.23: The dynamic loads on the bearings and the resulting fatigue between 12 s and 18 s. (a) and (b): the strategy 1. (c) and (d): the strategy 2. | 146 |
| Figure 8.24: The values of y sampled for the simulation study | 147 |
| Figure 8.25: The simulation results of the MCOS with the indexed y : Average accumulated fatigue | |

| | |
|--|-----|
| vs. Energy consumption..... | 148 |
| Figure 8.26: The simulation results of the MCOS with the indexed γ : Powertrain service life vs. Energy consumption..... | 149 |
| Figure 8.27: The simulation results of the MCOS with the indexed γ : Bearing accumulated fatigue vs. Energy consumption..... | 149 |
| Figure 8.28: The simulation results of the MCOS with the indexed γ : Bearing service life vs. Energy consumption..... | 150 |
| Figure 8.29: Gear distribution of the predictive MCOS with (a) $\gamma 1$, (b) $\gamma 31$ and (c) $\gamma 55$, respectively..... | 151 |
| Figure 8.30: Operating duration of both EMs with different γ : (a), (b) and (c): the EM1. (d), (e) and (f): the EM2..... | 152 |
| Figure 8.31: Load spectra of B1.1.2 (left side) and B2.1.3 (right side) with different γ | 153 |
| Figure 8.32: The simulation results of the MCOS with the indexed γ : Gear service life vs. Energy consumption..... | 154 |
| Figure 8.33: Shift counts of all driving cycle simulations..... | 155 |
| Figure 8.34: The simulation results of the MCOS with the indexed γ : Gear accumulated fatigue vs. Shift count. | 155 |
| Figure 8.35: Gear distribution of the MCOS with (a) $\gamma 43$, (b) $\gamma 45$ and (c) $\gamma 46$ | 156 |
| Figure 8.36: Stress spectra of G2.S1.P neg with different γ | 157 |
| Figure 8.37: Load spectra of B2.1.3 with different γ | 157 |
| Figure 8.38: Operating duration of the EM1 with different γ | 158 |
| | |
| Figure B. 1: Power loss maps of (a) the ST1, (b) the ST2 in the 1st gear position and (c) the ST2 in the 2nd gear position. | 170 |
| Figure B. 2: Power loss maps of (a) the EM1, (b) the LE1, (c) the EM2 and (d) the LE2. | 171 |
| | |
| Figure D. 1: Power loss at different angular velocities of (a) ST1, (b) EM1, (c) ST2 and (d) EM2 | 182 |
| Figure D. 2: Effective torques of the ST1 (a) and the ST2 in 1st gear (b) at different angular velocities | 183 |
| Figure D. 3: Torque range divided into several zones | 185 |
| | |
| Figure E. 1 Modified S-N Curves of tooth root breakage. (a) G1.S1.SG, (b) G1.S1.PG, (c) G1.S1.RG, (d) G1.S2.P, (e) G2.S1.P, (f) G2.S2.1P, (g) G2.S2.2P, (h) G2.S3.P and (i) G.FD.W. .. | 193 |
| Figure E. 2 Modified S-N Curves of tooth flank pitting. (a) G1.S1.SG, (b) G1.S1.PG, (c) G1.S1.RG, (d) G1.S2.P, (e) G2.S1.P, (f) G2.S2.1P, (g) G2.S2.2P, (h) G2.S3.P and (i) G.FD.W. .. | 194 |

| | |
|--|-----|
| Figure H. 1: Accumulated fatigue after a WLTC and corresponding service lives of the parts in Speed4E powertrain, controlled by the MCOS with γ 31. (a): Bearing fatigue. (b): Tooth root breakage. (c) Tooth flank pitting. | 198 |
| Figure H. 2: Accumulated fatigue after a WLTC and corresponding service lives of the parts in Speed4E powertrain, controlled by the MCOS with γ 55. (a): Bearing fatigue. (b): Tooth root breakage. (c) Tooth flank pitting. | 199 |
| Figure H. 3: The simulation results of the MCOS with the indexed γ : Gear accumulated fatigue vs. Energy consumption. | 200 |
| Figure H. 4: The values of γ sampled for the simulation study. | 200 |
| Figure H. 5: The simulation results of the MCOS with the indexed γ : Gear service life vs. Energy consumption. | 201 |
| Figure H. 6: The simulation results of the MCOS with the indexed γ : Bearing service life vs. Energy consumption. | 201 |

List of Tables

| | |
|---|-----|
| Table 4.1. Durations of different phases in a shift process | 49 |
| Table 4.2. Admissible discrete control set \mathcal{B} , controlled switching | 50 |
| Table 4.3. Admissible discrete control set \mathcal{B} , autonomous switching..... | 51 |
| Table 4.4. States in a shift process..... | 52 |
| Table 5.1: Variables influencing the time complexity of solveHMP | 76 |
| Table 5.2: Time complexity of each part in solveHMP | 77 |
| Table 7.1: Admissible discrete control set, controlled switching..... | 105 |
| Table 8.1 Preselection of parts for fatigue estimation..... | 118 |
| Table 8.2. An typical example of calculated $\mathbf{B1}$ and $\mathbf{B10}$ service life of A-parts in a transmission [153, p. 102] | 119 |
| Table 8.3. Parameters to calculate $\sigma\mathbf{H}$ | 126 |
| Table 8.4. Parameters to calculate \mathbf{N} | 127 |
| Table 8.5: The Gears in Speed4E powertrain..... | 129 |
| | |
| Table A. 1: Rotational inertia of the components in Speed4E powertrain | 168 |
| Table A. 2: Parameters of longitudinal dynamics..... | 169 |
| | |
| Table E. 1 Parameters to calculate $\sigma\mathbf{HG}$ | 190 |
| Table E. 2: Bearings used in Speed4E powertrain | 190 |
| Table E. 3: Parameters of the gears in Speed4E powertrain. | 191 |
| Table E. 4: Parameters of the bearings in Speed4E powertrain | 192 |
| | |
| Table G. 1 Values of the fatigue coefficients in MCOS with the fatigue cost functional for bearings. | 197 |

Nomenclature

Acronyms and Abbreviations

| Abbreviation | Description |
|---------------|-------------------------------------|
| 1-D (x -D) | One-dimensional (x -dimensional) |
| AMT | Automated manual transmission |
| APC | Angular position control |
| AT | Automatic transmission |
| BEV | Battery electric vehicle |
| DC | Dog clutch |
| DCT | Dual clutch transmission |
| DM | Direct method |
| DoF | Degree of freedom |
| DP | Dynamic programming |
| EFS | Externally forced switch |
| EM | Electric motor |
| EMS | Energy management strategy |
| EV | Electric vehicle |
| HEV | Hybrid electric vehicle |
| HMP | Hybrid minimum principle |
| HOCP | Hybrid optimal control problem |
| IFS | Internally forced switch |
| IM | Indirect method |
| IMS | Institute for Mechatronic Systems |
| LA | Linear actuator |
| LP | Linear programming |
| MC | Markov chain |
| MCOS | Multi-criteria operation strategy |
| MINLP | Mixed-integer nonlinear programming |
| MPBVP | Multi-point boundary value problem |
| NLP | Nonlinear programming |
| NN | Neural network |
| OCP | Optimal control problem |

| | |
|---------|--|
| PE | Power electronic |
| PHEV | Plug-in hybrid electric vehicle |
| PMP | Pontryagin Minimum Principle |
| RK | Runge-Kutta method |
| SDP | Stochastic dynamic programming |
| SOC | State of charge |
| Speed4E | Project “Speed4E” |
| SPT | Sub-powertrain |
| SQP | Sequential quadratic programming |
| SS | Shift sleeve |
| ST | Sub-transmission |
| TCU | Transmission control unit |
| TPM | Transition probability matrix |
| WLTC | Worldwide harmonized light vehicles test cycle |
| w.r.t. | With respect to |

Roman Symbols

| Symbol | Description |
|---------------|---|
| A | State-jump function |
| A_f | Vehicle frontal area |
| \mathcal{A} | Set of state-jump functions |
| B_x | B_x -service life |
| \mathcal{B} | Admissible control set |
| \mathcal{B} | Set of admissible control sets |
| c_{aero} | Aerodynamic resistance coefficient |
| c_B | Root of torque coefficient for bearing dynamic loads |
| c_G | Torque coefficient for effective bending stress on root areas |
| c_H | Torque coefficient for Hertzian stress on tooth flanks |
| c_{roll} | Rolling resistance coefficient |
| \mathcal{C} | Collection of constraint functions |
| d | Fatigue rate |

| | |
|----------------------|---|
| D | Fatigue |
| $Dist$ | Distance |
| \mathbf{f} | Vector field |
| F_a, F_n, F_r, F_t | Axial, normal, radial and transverse decompositions of meshing forces |
| F_{axial} | Axial load of a bearing |
| F_{radial} | Radial load of a bearing |
| \mathcal{F} | Collection of indexed vector fields |
| g | Gravity acceleration |
| G | Gear position |
| \mathcal{G} | Discretization grid |
| h | Time step size |
| \mathcal{H} | Hamiltonian |
| \mathbb{H} | Hybrid system |
| i | Gear ratio |
| I_{red} | Reduced rotational inertia |
| J | Cost functional |
| l | Cost function |
| l_{switch} | Switching cost function |
| L | Equivalent dynamic load |
| m | Endpoint functional |
| $mass$ | Vehicle mass |
| mc | Markov chain state |
| M | Switching manifold |
| \mathcal{M} | Set of switching manifolds |
| M | Number of motors in an electric powertrain |
| N | Permissible cycle |
| N_{exe} | Switching execution count |
| N_c | Length of control horizon |
| N_D | Permissible cycle at fatigue limit |
| N_p | Length of prediction horizon |
| N_Q | Size of a discrete set Q |
| N_t | Length of time horizon |
| N_u | Dimension of continuous controls \mathbf{u} |

| | |
|---------------------|---|
| N_x | Dimension of continuous states \mathbf{x} |
| N | Number of multi-speed transmissions in an electric powertrain |
| \mathbb{N} | Natural number set |
| O | Big-O notation for the worst case time complexity |
| P | Power |
| PL | Power loss |
| q | Discrete state |
| $q(\cdot)$ | Discrete state sequence |
| Q | Finite set of discrete states of a hybrid system |
| r | Wheel dynamic radius |
| \mathbb{R}^n | n -dimensional real number set |
| SL | Service life |
| t | Time |
| t_0 | Initial time |
| t_f | Final time |
| tpm | Element in transition probability matrices |
| T | Torque |
| T_{in} | Input torque transmitted by a gearset |
| T_{total} | Output torque of a powertrain |
| TS | Change rate of the output torque of a sub-transmission |
| t | Hybrid time trajectory of switching |
| \mathbf{u} | Continuous controls |
| $\mathbf{u}(\cdot)$ | Continuous control function |
| \mathcal{U} | Admissible continuous control space |
| \mathbb{U} | Continuous-valued control space |
| v | Vehicle speed |
| \mathbf{V} | Sequence of vehicle speed |
| $\dot{\mathbf{V}}$ | Sequence of acceleration |
| \mathbf{x} | Continuous states |
| $\mathbf{x}(\cdot)$ | Continuous state function |
| \mathcal{X} | Admissible continuous state space |
| \mathbb{X} | Continuous-valued state space |
| \mathbf{y} | Extended states |

\emptyset empty set

Greek Symbols

| Symbol | Description |
|----------------------|--|
| α | elektromagnetische Luftspaltwellenausbreitungskonstante |
| β | Coefficient of an endpoint functional |
| γ | Coefficient of a cost functional |
| Γ | Dynamics of extended states |
| δ | Change of value |
| ϵ | Tolerance of the first order optimality condition |
| ε | Convergence tolerance |
| λ | Costates |
| $\lambda(\cdot)$ | Costate function |
| π | Combination of discrete states |
| ϖ | Discrete control |
| $\varpi(\cdot)$ | Discrete control sequence |
| Π | Discrete transition function |
| Π | Discrete set of roman numbers |
| ρ_{air} | Air density |
| σ | Stress |
| σ_{F} | Effective root bending stress |
| σ_{H} | Hertzian stress |
| τ | Duration |
| ϕ | Transmission effective torque function |
| ϕ' | Function calculating input torques based on transmission effective torques |
| ϕ_{ω} | Transmission effective torque function with a fixed input angular velocity |
| ϕ_{ω}^{-1} | Inverse function of ϕ_{ω} |
| Φ | Error of transversality condition |
| ψ | Boundary condition |
| ω | Angular velocity |
| Ω | Sample space |

Superscripts

| Index | Description |
|-------|------------------------------|
| + | Right after a time instance |
| – | Right before a time instance |
| * | Optimal |
| T | Transpose |

Subscripts

| Index | Description |
|-----------|---|
| + | Right after a time instance |
| – | Right before a time instance |
| <i>a</i> | “Balance” and “Decrease” phases in Speed4E shift processes |
| acc | Acceleration of an electric motor |
| <i>b</i> | The first “Synchronize” and “Disengage” phases in Speed4E shift processes |
| B | Bearing |
| Bal | “Balance” phases Speed4E shift processes |
| <i>c</i> | The second “Synchronize” and “Engage” phases in Speed4E shift processes |
| cycle | Driving cycle |
| <i>d</i> | “Increase” and the second “Balance” phases in Speed4E shift processes |
| Decr | “Decrease” phase in Speed4E shift processes |
| Disengage | “Disengage” phase in Speed4E shift processes |
| current | Current gear |
| ele | Electric |
| Engage | “Disengage” phase in Speed4E shift processes |
| eq | Equivalent |
| Fade | Torque of an electric motor fades to a new value in “Balance” phases |
| G | Gear |
| Incr | “Increase” phase in Speed4E shift processes |
| max | Maximal |
| min | Minimal |
| multi | Multi-criteria |
| neg | Tooth flank in contact, when a gearset transmits a negative torque |
| pos | Tooth flank in contact, when a gearset transmits a positive torque |

| | |
|------------|---|
| p, pred | Predictive |
| req | Driver requested |
| ST2, input | Rotational inertia around |
| syn1, syn2 | The First and the second "Synchronize" phases |
| target | Target gear |

1 Introduction

Within the scope of “The European Green Deal” [1], the European Union (EU) strives to be the first climate-neutral¹ continent without net emissions of greenhouse gases by 2050. Electric vehicles (EV) are the key technology to decarbonise road transport, a sector that accounts for 16 % of global emissions [3]. The worldwide EV market, including Battery Electric Vehicles (BEV) and Plug-in Hybrid Electric Vehicles (PHEV), has witnessed an exponential growth, as their sales exceeded 10 million in 2022, while the share of EV in the total vehicle sales has tripled from 4 % in 2020 to 14 % in 2022 [4]. Figure 1.1(a) illustrates the annual sales of EV from 2016 to 2023 in the world. The EV Sales in Europe have grown steadily to 2.7 million in 2022, and are estimated to reach 3.4 million in 2023².

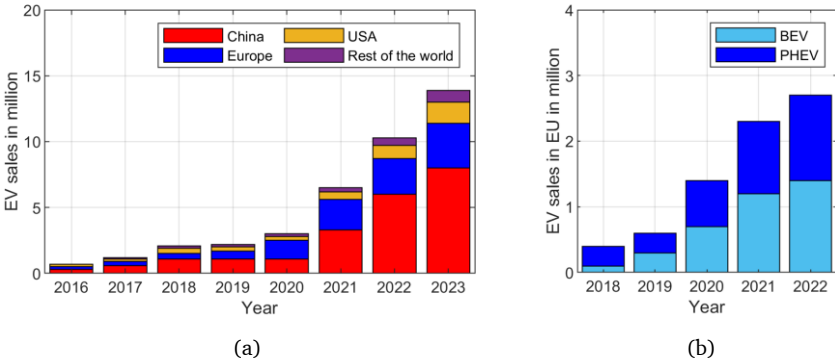


Figure 1.1: EV sales volumes [4]. (a) EV worldwide sales from 2016 to 2023³. (b) BEV and PHEV sales in EU from 2018 to 2022.

A more progressive regulation, the revised rules on CO₂-emission performance standards [5], entered into force under the agreement of EU member states in March, 2023. Under the new regulation, from 2035, all new cars and vans registered in the EU are set to be zero emission. In the current market, BEVs are the one of the main solutions to the regulation. Figure 1.1(b) illustrates the BEV sales in the EU, which has risen steadily from 0.1 million in 2018 to 1.4 million in 2022. BEVs took up more than 10 % of total vehicle sales in 2022. If the regulation stays in force and

¹ Climate neutrality refers to the idea of achieving net zero greenhouse gas emissions by balancing those emissions so they are equal (or less than) the emissions that get removed through the planet’s natural absorption [2].

² At the time of writing, the EV sales in 2023 are not yet available from the International Energy Association.

³ 2023 sales are estimated based on market trends through the first quarter of 2023.

major breakthroughs in other technologies are absent, BEV sales will soar in the near future.

Range of BEVs was considered one of the main drawbacks that keep consumers from choosing BEVs, claimed a study conducted in 2015 [6]. Though further developed since then⁴, range of BEVs still falls short of conventional vehicles. It is essential to further extend range of BEVs so that they can replace conventional vehicles as main cars and vans newly registered in the EU under the new regulation from 2035.

1.1 Motivation

Additional to increasing battery capacity, reducing energy consumption helps to increase range of BEVs. It was reported in multiple works [7–10] that powertrains with multi-speed transmissions can reduce energy consumption of BEVs. It was shown in recent studies [11–13] that, under the same requirements of dynamic performance, two-drive powertrains, i.e. powertrains consisting of two electric motors (EM), consume considerable less energy in Worldwide Harmonized Light Vehicles Test Cycle (WLTC) than their single-drive counterparts. Powertrains with either concept gain an extra degree of freedom (DoF), thanks to which operating points of components in powertrains can vary in favour of a lower total energy consumption, either through changing gear position or through splitting total output power differently between both EMs.

Necessity of energy management strategies

To combine the benefits of both concepts, two-drive multi-speed powertrains are proposed for BEVs. A control strategy is essential for such powertrains with more than one DoF to fulfil basic requirements, e.g. requested acceleration, comfort, and most importantly to achieve minimal energy consumption. Such a control strategy is called an *Energy Management Strategy* (EMS)⁵. Figure 1.2 shows a typical control scheme of a BEV powertrain, in which the EMS functions as a high-level control that determines control set points for low-level controls, i.e. the EM torque (Trq) for the MCU and the target gear position (G_{target}) for the TCU, based on the driver requested power (P_{req}), the vehicle speed (v_{veh}), the EM rotational velocity (ω_{EM}), the current gear position ($G_{current}$) and the state of charge (SOC). In general, low-level controls regulate the components in a powertrain based on the set points provided by a high-level control.

⁴ For instance, the Worldwide Harmonized Light Vehicles Test Procedure range of BMW i3 has increased from 235 km (model year 2018) to 308 km (model year 2022), largely thanks to the battery capacity increasing from 94 Ah (model year 2018) to 120 Ah (model year 2022).

⁵ It is also referred to as supervisory control strategy in the literature, e.g. [14] and [15].

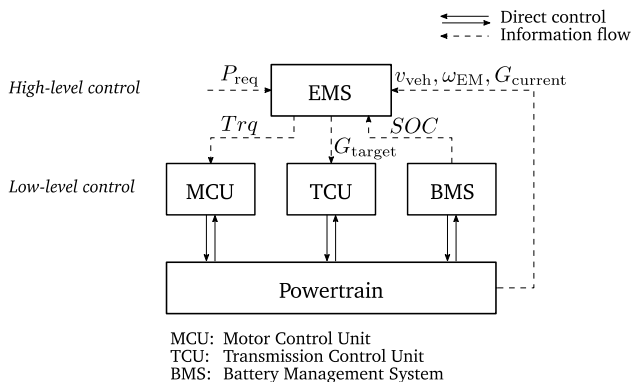


Figure 1.2: EMS as high-level control in a typical BEV powertrain control scheme

Observations of relevant energy management strategies in the literature

To design an EMS that minimizes energy consumption is to answer the question: given a mathematical model of a powertrain and a vehicle as well as constraints of components, *how can their speed, torque and gear position be controlled so that minimal energy consumption and other additional performance criteria can be achieved?* Such a problem falls into the category of *Optimal Control Problem (OCP)*, which is to determine the control function that causes a process to satisfy the physical constraints and at the same time minimize (or maximize) some performance criteria [16, p. 3]. More specifically, due to the presence of G_{target} and G_{current} , a discrete control and a discrete state, a two-drive multi-speed powertrain is naturally a *Hybrid System*⁶, a system that has continuous controls, e.g. EM torques, but at the same time makes discrete decisions, e.g. target gears, instead of a pure continuous system. Consequently, the OCP in question is a *Hybrid Optimal Control Problem (HOCP)*⁷ [19, 20].

There are numerous studies on EMSs that consider powertrains with more than one power source⁸ and multi-speed transmissions. However, such powertrains were mostly not explicitly

⁶ A continuous system that can switch between different subsystems, i.e. differential equations describing the time derivative of continuous states, can also be modelled as a hybrid system, e.g. *A bouncing ball* in Section 2.1.1.

⁷ In the optimization community, it is more commonly referred to as mixed-logic dynamic optimization [17] or mixed-integer optimal control problem [18]. Such terminologies highlight the mixed-integer optimization characteristic in the problem.

⁸ Machines designed to convert other types of energy, e.g. chemical energy and electric energy, to mechanical energy.

modelled as hybrid systems in the literature, for example [14, 21–23] to name but a few. The resulting EMSs were often treated with solution methods for continuous OCPs. Such treatment does not necessarily indicate flawed results. For instance, results principally derived from the discrete *Dynamic Programming* (DP), e.g. [24], [25] and [21], hold, since the method is inherently able to deal with systems involving discrete variables. Nevertheless, proper system formulation and problem solution are desirable.

For the goal of real-world applications, it is necessary to design an EMS that does not consider speed profiles *a priori*. Such an EMS is called an online EMS. Furthermore, due to the limited computational resources in real-world applications, it is essential to design an EMS achieving high performance, i.e. the controlled vehicle consumes as little as possible energy in evaluations, and requiring short computational time⁹.

As far as the author knows, other than [26], EMSs of powertrains with multi-speed transmissions were developed in the literature with an assumption: “shift processes are ignored”. The most common transmission types in today’s market are automatic transmissions (AT), dual clutch transmissions (DCT) and automated manual transmissions (AMT). Shift durations¹⁰ of some examples from these types of transmission are 0.8 s – 1.6 s [27, 28], 0.4 s – 0.7 s [29, 30], and 0.7 s – 1.3 s [31, 32], respectively. Though not covering all models from the mentioned transmission types, the statistics show that shift durations of ATs and AMTs can exceed 1 s, while shift durations of DCTs are smaller than but not distant from 1 s. It might not be most rigorous to develop EMSs with a frequency of at least 1 Hz ignoring shift processes for powertrains with above mentioned transmission types in [24, 25, 33–35], to name but a few. Though the EMS developed in [26] considered shift processes, it was based on a predefined shift sequence in a time span of ca. 6 s for a single-drive BEV, whose results lacked association to energy consumption. An EMS with proper consideration of shift processes is valuable but absent.

Today’s EV market has witnessed not only two-drive BEVs but also three-drive BEVs, e.g. the Tesla Model S plaid, the Tesla Cybertruck Cyberbeast and the Lucid Air, and four-drive BEVs, e.g. the Zeekr 001 FR. In the future, vehicles—not necessarily passenger cars but ground vehicles, heavy-duty vehicles, water vehicles etc.—might be equipped with electric powertrains with even more EMSs for their specific purposes. Is there a general way to design their EMSs? If so, how? The questions are ambitious. It might be a good idea to start with the EMS of a two-drive BEV and try to extend the results to a general multi-drive multi-speed BEV. Specifically, the general multi-drive powertrain has M EMSs and transmissions parallelly coupled, $M \geq 2$, $M \in \mathbb{N}$.

⁹ Evaluation of computational time will be more technically discussed in Section 2.4.4.

¹⁰ Shift duration is a period between the moment when a shift is signalled and the synchronization of transmission input speed into its new value according to target gears.

Suppose that an EMS is designed. Under its control, the parts and the components in a powertrain are loaded—some more than others¹¹—in favour of lower total energy consumption. One can imagine, these more loaded parts, e.g. gears, would reach their service lives earlier than other same kind of parts, if they would be designed to reach the same service life under the same loads. Such an outcome limits the service life of the powertrain, since the reliability and the service life of a powertrain is dependent on those of its parts and components [36, p. 251]. Though this is often addressed in design processes [36, p. 279], it is beneficial to design an EMS considers service life aspect, since it is the EMS that controls a powertrain on a high-level and determines loads of its parts to an extent. It is noteworthy that such an attempt is unseen in the literature other than in the previous work of the author [37].

1.2 Goals and Contributions

Firstly, a proper mathematical model is to be formulated for a two-drive multi-speed BEV in such a way that captures its continuous and discrete dynamics, and the shift processes are properly considered.

Secondly, an online EMS is to be designed. A proper accurate solution method is to be implemented, whose algorithm operates with lower time complexity compared to its peers. The algorithms of the EMS are to be presented.

Thirdly, the solution of the HOCP and the developed algorithm are studied from a mathematical point of view with the goal of reducing time complexity.

As far as the author knows, a hybrid system formulation of a two-drive multi-speed powertrain is missing in the literature. Furthermore, an EMS that considers shift processes and incorporate proper numerical methods was not reported in the literature. More importantly, reducing the time complexity of a solution method through mathematical analysis provides great value and can be transferred to other problems with a similar setting.

The system formulation and the solution of the HOCP shall be able to be extended to accommodate multi-drive multi-speed electric powertrains. For this purpose, a general multi-drive multi-speed electric powertrain is proposed with the general hybrid system and the general HOCP. The solution method and the algorithm for the online EMS are developed in such a way that can be applied to the general HOCP. As a whole, a unified framework to design an EMS for the general multi-drive multi-speed BEV is to be obtained, which is unseen in the literature.

Lastly, a multi-criteria operation strategy considering both energy consumption and service life

¹¹ See Section 8.1 for a more systematic discussion of loads.

of a powertrain is to be developed. Such a concept is only seen in the previous work of the author [37]. This thesis addresses this topic in a more systematic way, thanks to the consideration of shift processes and a more inclusive service life model.

1.3 Structure of the Content

This thesis is organized as follows:

In Chapter 2, background knowledge is provided. Hybrid systems are introduced with a simple example. Formulation of a HOCP is shown, so that the meaning of “solving a HOCP” is clear. Thereafter, solution methods for HOCPs are reviewed in Section 2.4 regarding their main principles, which will be supplemented with simple examples. Solution methods are compared against each other w.r.t. relevant properties to online EMS.

EMSs developed in the literature will be reviewed in Chapter 3, where, among other aspects, the treatment of continuous and discrete variables will be discussed. A proper solution method for the online EMS is to be identified after the review.

In Chapter 4, a two-drive multi-speed powertrain is modelled as a hybrid system, which provides a specific model for the formulation of the HOCP in this chapter and the EMS in the later chapters.

Based on the mathematical model, an online EMS is developed in Chapter 5. The algorithms including the numerical methods are presented. The optimal control problem is evaluated. The time complexities of the developed algorithms are analysed.

In Chapter 6, the online EMS is evaluated from two perspectives: 1) individual solution processes and 2) driving cycle simulations.

A general multi-drive multi-speed powertrain is presented in Chapter 7. It is showcased, how to modify several small parts of the general hybrid system to achieve different operation complexity. The general HOCP is presented and analysed. A unified framework to design an EMS is presented.

In Chapter 8, the online EMS is systematically further developed to incorporate the view point of powertrain service life, whose results are reported.

Conclusion and outlooks are presented in Chapter 9.

1.4 Acronyms and Abbreviations

There are plenty of acronyms and abbreviations in this thesis. To avoid confusion, this subsection clarifies, how they are categorized and treated in the writing.

“EM”, “EMS”, etc. are considered initialisms. They take the definite article and their plural forms

end with an “s”. In some cases, they are indexed, e.g. “EM1”, but still considered initialisms.

A powertrain developed in a project called “Speed4E” will be mentioned repeatedly. Speed4E is considered an acronym and, therefore, does not take the definite article.

In Chapter 8, the bearings and gears in Speed4E powertrain are labelled. For instance, a bearing is labelled “B1.1.1”, and it is considered an acronym.

2 Background Knowledge

This chapter provides the necessary background knowledge about hybrid systems and HOCs. The first part of this chapter provides the definition of hybrid systems. The second part formulates HOCs. Discretization and integration schemes are discussed in the third part, which are fundamental for numerical solutions. The fourth part reviews the mathematical principles that are used to solve HOCs with abstract but illustrative examples. Their properties are outlined to provide arguments for selecting one of them to be implemented in this work.

2.1 Hybrid Systems and Regularity Conditions

The first subsection gives the definition of hybrid systems with a simple example. A hybrid system must fulfil certain conditions, so that an optimal control for a HOC regarding the hybrid system exists. These conditions are called regularity conditions. They are provided with little discussion in the second subsection. The readers that are interested in the existence and uniqueness of optimal controls can refer to [38–40].

2.1.1 Hybrid Systems

A hybrid system consists of both continuous and discrete states. It is formally defined by:

Definition 2.1 *Hybrid System* [41, p. 14]: a general hybrid system is a 9-tuple

$$\mathbb{H} := (\mathcal{Q}, \mathbb{X}, \mathbb{U}, \mathcal{F}, \mathcal{A}, \mathcal{B}, \Pi, \mathcal{M}, \mathcal{C}). \quad (2.1)$$

\mathcal{Q} denotes a finite set of discrete states. The discrete state of the hybrid system $q \in \mathcal{Q}$ represents a subsystem of the hybrid system \mathbb{H} . $\mathbb{X} := \{\mathbb{X}_q \subseteq \mathbb{R}^{N_x}\}$ is a collection of continuous-valued state spaces, where \mathbf{x} is a N_x -dimensional continuous-valued state vector. It's called continuous states in this thesis. $\mathbb{U} := \{\mathbb{U}_q \subseteq \mathbb{R}^{N_u}\}$ is a collection of continuous-valued control spaces for each discrete state. \mathbf{u} is a N_u -dimensional continuous-valued control vector. It's called continuous controls in this thesis. For each $q \in \mathcal{Q}$, one vector field $\mathbf{f}_q: \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{X}$ is defined. \mathcal{F} denotes an indexed collection of them $\mathcal{F} := \{\mathbf{f}_q: \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{X} \mid q \in \mathcal{Q}\}$. $\mathcal{A} := \{A_{q_i, q_j} \mid q_i, q_j \in \hat{\mathcal{Q}}, q_i \neq q_j\}$ is a set of state-jump functions. $\Pi: \mathbb{X} \times \mathcal{Q} \times \Pi_q \rightarrow \mathcal{Q}$ is a discrete transition function. $\mathcal{B} := \{\mathcal{B}_q \mid q \in \mathcal{Q}\}$ is a collection of discrete sets $\mathcal{B}_q \subseteq \mathcal{Q}$, in which admissible discrete controls edges $\varpi_q \in \mathcal{B}_q$ are defined. $\mathcal{M} := \{M_{q_i, q_j} \mid q_i, q_j \in \hat{\mathcal{Q}}, q_i \neq q_j\}$ denotes a set of time-dependent switching manifolds, which determines autonomous switching. \mathcal{C} is a collection of constraint functions. It includes mixed control-state constraints, pure state constraints and pure control constraints.

△

A hybrid system can be interpreted as a collection of N_q number of subsystems, which are indexed by the discrete state and whose dynamics are determined by the respective vector field. There are a continuous state space $\mathbb{X}_q \subset \mathbb{R}^{N_x}$ and a continuous control space $\mathbb{U}_q \subset \mathbb{R}^{N_u}$ for each subsystem. In this thesis, the dimension of the continuous state space and that of the continuous control space are assumed constant. The continuous states and the continuous controls are constrained by the constraints from \mathbf{C} , which forms an admissible continuous state space \mathcal{X} and an admissible continuous control space \mathcal{U} . They are discrete state dependent, i.e. $\mathbf{x} \in \mathcal{X}_q$ and $\mathbf{u} \in \mathcal{U}_q$, $q \in \mathcal{Q}$, if the constraints are discrete state dependent.

The dynamics of the hybrid system are switched among the subsystems on the time interval $[t_0, t_f]$, where t_0 and t_f denote the initial and final times. In this thesis, they are considered fixed and finite. There are in total $0 \leq N_{\text{exe}} < \infty$ number of switching happened in the time interval, which form a *discrete control function* $\varpi(\cdot) = (\varpi_1, \varpi_2, \dots, \varpi_j, \dots, \varpi_{N_{\text{exe}}})$. The time instances that the switching happen form a strictly increasing sequence $\mathbf{t} = (t_1, t_2, \dots, t_j, \dots, t_{N_{\text{exe}}}), t_j \in [t_0, t_f]$. Based on the discrete control, the discrete state, and the continuous states at a time instance, the discrete transition function Π determines the coming discrete state,

$$q(t_j^+) = \Pi(\mathbf{x}(t_j^-), q(t_j^-), \varpi_j), \quad (2.2)$$

where the minus and the plus signs in the superscripts of the time instances refer to the right before and right after the time instances. Four types of switching can happen: *autonomous switching*, *autonomous jumps*, *controlled switching*, and *controlled jumps* [42].

A transition is named a “*switching*” if the state-jump function $A_{q_{t_j^-}, q_{t_j^+}}$ does not exist or it produces a zero when the transition in (2.2) happens. Otherwise, a “*jump*” happens, which causes the continuous states to exhibit discontinuity at the time instance of the transition. “*Autonomous*” indicates that the discrete state at t_j changes in a predefined way that is determined by the transition function Π , when the continuous states meet a switching manifold $M_{q_{t_j^-}, q_{t_j^+}}$. It is called internally forced switching in [43]. “*Controlled*”, on the other hand, indicates that the discrete control is consciously chosen. It is called externally forced switching in [43]. From the point of view from control, autonomous switching and jumps are completely described by the transition function Π , whereas controlled switching and jumps need the discrete control ϖ to initiate the transition.

Most of the real systems operate with physical constraints on the continuous controls and the continuous states. The set \mathbf{C} in \mathbb{H} by **Definition 2.1** may possess different types of constraints:

Definition 2.2 *Mixed constraints, State constraints, and Control constraints:* The *mixed constraints* depend on both \mathbf{x} and \mathbf{u} at the same time,

$$c_{\text{mix},i}(\mathbf{x}(t), \mathbf{u}(t), q(t)) \leq 0, \quad i = 1, \dots, m, \quad \mathbf{x}(t) \in \mathbb{X}, \quad \mathbf{u}(t) \in \mathbb{U}, \quad q \in \mathcal{Q}, \quad \forall t \in [t_0, t_f], \quad (2.3)$$

where i indexes the constraints and m is the total number of the constraints. For compactness, (2.3) is rewritten as

$$\mathbf{c}_{\text{mix}}(\mathbf{x}(t), \mathbf{u}(t), q(t)) \leq \mathbf{0}, \quad \mathbf{x}(t) \in \mathbb{X}, \quad q \in \mathcal{Q}, \quad \forall t \in [t_0, t_f], \quad (2.4)$$

where $\mathbf{c}_{\text{state}}: \mathbb{X} \times \mathcal{Q} \rightarrow \mathbb{R}^m$ produces a m -dimensional vector and $\mathbf{0}$ is a m -dimensional zero vector. The inequality “ \leq ” is evaluated elementwise.

State constraints: The *state constraints* depend on \mathbf{x} , i.e.,

$$\mathbf{c}_{\text{state}}(\mathbf{x}(t), q(t)) \leq \mathbf{0}, \quad \mathbf{x}(t) \in \mathbb{X}, \quad q \in \mathcal{Q}, \quad \forall t \in [t_0, t_f]. \quad (2.5)$$

Control constraints: The *control constraints* depend on \mathbf{u} , i.e.,

$$\mathbf{c}_{\text{control}}(\mathbf{u}(t), q(t)) \leq \mathbf{0} \quad \mathbf{u}(t) \in \mathbb{U}, \quad q \in \mathcal{Q}, \quad \forall t \in [t_0, t_f]. \quad (2.6)$$

△

Under the constraints from the set \mathcal{C} , \mathbf{x} and \mathbf{u} of the system \mathbb{H} by **Definition 2.1** are restricted in certain spaces, which are called *admissible state sets* and *admissible control sets*.

Definition 2.3 *Admissible State Sets and Admissible Control Sets:* the admissible state sets that depend on the constraints given by (2.4) and (2.5) are defined by

$$\mathcal{X}(\mathbf{u}(t), q(t)) := \{\mathbf{x}(t) \in \mathbb{X} \mid \mathbf{c}_{\text{mix}}(\mathbf{x}(t), \mathbf{u}(t), q(t)) \leq \mathbf{0}, \mathbf{c}_{\text{state}}(\mathbf{x}(t), q(t)) \leq \mathbf{0}\}. \quad (2.7)$$

The admissible control sets depending on the constraints (2.4) and (2.6) are defined by

$$\mathcal{U}(\mathbf{x}(t), q(t)) := \{\mathbf{u}(t) \in \mathbb{U} \mid \mathbf{c}_{\text{mix}}(\mathbf{x}(t), \mathbf{u}(t), q(t)) \leq \mathbf{0}, \mathbf{c}_{\text{control}}(\mathbf{u}(t), q(t)) \leq \mathbf{0}\}. \quad (2.8)$$

△

In this thesis, Zeno-behaviours, i.e. infinite switching in finite time, are excluded.

It is necessary to elaborate how a hybrid system evolves, so that the definition above is understandable. Imagine a dimensionless small ball released at the height of h_0 and it bounces at the floor, as shown in Figure 2.1. The continuous states of the ball are its height and the velocity $\mathbf{x} =$

$(h, \dot{h})^\top$. To have the modelling process simplified, the contact between the ball and the floor is considered instant, at which a constant proportion $1 - 2\gamma^2$ of the kinetic energy is lost.

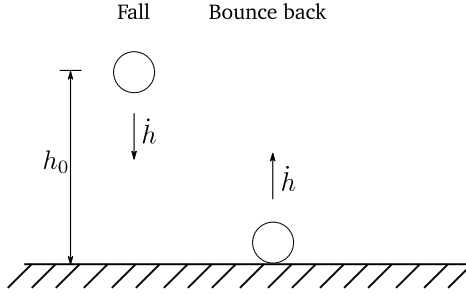


Figure 2.1: A bouncing ball

The bouncing ball can be modelled as a hybrid system with $\mathcal{Q} := \{\text{"Fall"}, \text{"Bounce back"}\}$, $\mathbb{X} := \mathbb{R}^2$, $\mathbb{U} := \emptyset$. The vector fields of both discrete states are determined by the gravity acceleration g

$$\dot{\mathbf{x}} = \mathbf{f}_{\text{Fall}} = \begin{pmatrix} \dot{h} \\ -g \end{pmatrix}, \quad (2.9)$$

$$\dot{\mathbf{x}} = \mathbf{f}_{\text{Bounce back}} = \begin{pmatrix} \dot{h} \\ -g \end{pmatrix}. \quad (2.10)$$

The discrete control $\varpi(t_j) = \text{"Bounce back"}$ is chosen, when the discrete state is "Fall" and the manifold $M_{\text{FB}}(h) = 0$ is hit. The discrete transition function Π determines the system to switch to "Bounce back"

$$q(t_j^+) = \Pi(\mathbf{x}, q(t_j^-), \varpi(t_j)) = \text{"Bounce back"}, \quad (2.11)$$

where t_j denotes the time point, when a switch takes place. At the instant of the transition from "Fall" to "Bounce back", the state-jump function A_{FB} determines the continuous states at the right side of the time point based on the energy loss characterized by "coefficient"

$$\mathbf{x}(t_j^+) = A_{\text{FB}} = \begin{pmatrix} 1 & 0 \\ -1 & \text{coefficient} \end{pmatrix} \cdot \mathbf{x}(t_j^-). \quad (2.12)$$

The transition is categorized as an autonomous jump.

Likewise, the discrete control $\varpi(t_j) = \text{"Fall"}$ is chosen, when the discrete state is "Bounce back" and the manifold $M_{\text{BF}}(\dot{h}) = 0$ is hit. The discrete transition function Π determines the system to

switch to “Fall”

$$q(t_{j'}^+) = \Pi \left(\mathbf{x}, q(t_{j'}^-), \varpi(t_{j'}) \right) = \text{“Fall”}, \quad (2.13)$$

where $t_{j'}$ is used to differ from t_j . The state-jump function A_{BF} is an identity matrix of size 2. The transition is categorized as an autonomous switching.

The continuous states can be constrained by the discrete-state-dependent linear inequalities as

$$\begin{aligned} \mathcal{X}_{\text{Fall}} &:= \left\{ \mathbf{x} \mid \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \mathbf{x} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}, \\ \mathcal{X}_{\text{Bounce back}} &:= \left\{ \mathbf{x} \mid \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \mathbf{x} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}, \end{aligned} \quad (2.14)$$

which form the collection $\mathcal{C} := \{\mathcal{X}_{\text{Fall}}, \mathcal{X}_{\text{Bounce back}}\}$. The rest of the collections in **Definition 2.1** are $\mathcal{F} := \{f_{\text{Fall}}, f_{\text{Bounce back}}\}$, $\mathcal{A} := \{A_{FB}, A_{BF}\}$, $\mathcal{B} := \{\mathcal{B}_{\text{Fall}} := \{\text{Fall, Bounce back}\}, \mathcal{B}_{\text{Bounce back}} := \{\text{Fall, Bounce back}\}\}$ and $\mathcal{M} := \{M_{FB}, M_{BF}\}$. Though identical, the discrete sets $\mathcal{B}_{\text{Fall}}$ and $\mathcal{B}_{\text{Bounce back}}$ are explicitly defined for clarity. The hybrid system can be illustrated as a hybrid automaton, as shown in Figure 2.2, in which the autonomous transitions, either switching or jumps, are marked with red arrowed curves. Controlled transitions, either switching or jumps, are marked with blue arrowed curves in the hybrid automata in this work.

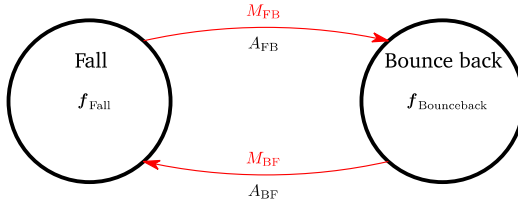


Figure 2.2: Hybrid automaton of a bouncing ball

Since the vector fields of both discrete states are identical and the state-jump function A_{BF} is an identity matrix, the hybrid system can be reduced to a system of one discrete state with modification of the above defined hybrid system with two discrete states.

What is of particular interest for this work are hybrid systems that exhibit no state-jump at switching. Such systems can be obtained from **Definition 2.1** by setting the set of state-jump function \mathcal{A} in **Definition 2.1** empty. The definition of a hybrid system without state-jump is obtained as follows:

Definition 2.4 *Hybrid System without State-Jump*: a hybrid system without state-jump consists of an 8-tuple

$$\mathbb{H}_1 := (\mathcal{Q}, \mathbb{X}, \mathbb{U}, \mathcal{F}, \mathcal{B}, \Pi, \mathcal{M}, \mathcal{C}). \quad (2.15)$$

△

2.1.2 Regularity Conditions

A prerequisite for the existence of an optimal control is the uniqueness of the trajectories of hybrid systems, i.e. the trajectory pair $(\mathbf{x}(\cdot), q(\cdot))$, see **Definition 2.5** and **Definition 2.7**, is unique on the time interval $[t_0, t_f]$ for every choice of the initial continuous and discrete states $(\mathbf{x}(t_0), q(t_0))$ as well as every admissible continuous and discrete control functions $(\mathbf{u}(\cdot), \varpi(\cdot))$. This subsection firstly lays out the relevant concepts and secondly provide the assumptions, based on which the uniqueness holds.

Definition 2.5 *Continuous state trajectory*: The continuous states \mathbf{x} on the time interval $[t_0, t_f]$ forms a *Continuous state trajectory* $\mathbf{x}(\cdot)$, which can be considered a function from $[t_0, t_f]$ to \mathbb{X} . It is also called a *solution* in other literatures. The continuous states are bounded by the constraints from the set \mathcal{C} , which forms an admissible function space $\mathcal{X}(\mathbf{u}(\cdot), q(\cdot))$ for the continuous state trajectory $\mathbf{x}(\cdot)$.

△

Definition 2.6 *Continuous control function*: The continuous controls \mathbf{u} on the time interval $[t_0, t_f]$ forms a *continuous control function* $\mathbf{u}(\cdot)$, which can be considered a function from $[t_0, t_f]$ to \mathbb{U} . The continuous controls are bounded by the constraints from the set \mathcal{C} , which forms an admissible function space $\mathcal{U}(\mathbf{x}(\cdot), q(\cdot))$ for the continuous control function $\mathbf{u}(\cdot)$.

△

Definition 2.7 *Discrete control function and discrete state trajectory*: A *hybrid time trajectory* is a strictly increasing sequence $\mathbf{t} = (t_0, t_1, \dots, t_j, \dots, t_{N_{\text{exe}}})$ with $t_j \in [t_0, t_f]$ and $t_{N_{\text{exe}}} = t_f$, where $0 \leq N_{\text{exe}} < \infty$ denotes the amount of switching in the time interval. The discrete control ϖ on the time interval $[t_0, t_f]$ forms *discrete control functions* $\varpi(\cdot) = (\varpi_0, \varpi_1, \dots, \varpi_j, \dots, \varpi_{N_{\text{exe}}-1})$ with piecewise functions $\varpi_j: [t_j, t_{j+1}) \rightarrow \mathcal{Q}$. The discrete state q on the time interval $[t_0, t_f]$ forms a *discrete control trajectory* $q(\cdot) = (q_0, q_1, \dots, q_j, \dots, q_{N_{\text{exe}}-1})$ with piecewise functions $q_j: [t_j, t_{j+1}) \rightarrow \mathcal{Q}$. The discrete control is bounded by the discrete set Π depending on q , which imposes an admissible function space $\mathcal{B}_{q(\cdot)}$ on the discrete control function $\varpi(\cdot)$. The admissible function space of the discrete state trajectory $q(\cdot)$ is $\mathcal{Q}(\cdot)$.

The assumptions of the hybrid system \mathbb{H} to ensure the uniqueness are listed as following, which are considered to hold in the rest of the work.

Assumption 2.1: All continuous control spaces \mathbb{U}_q are compact sets, i.e. they are closed and bounded¹². The continuous control function $\mathbf{u}(\cdot): [t_0, t_f] \rightarrow \mathbb{U}$ is measurable¹³.

Assumption 2.2: All vector fields $f_q: \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{X}$ are at least once continuously differentiable w.r.t. the continuous states \mathbf{x} and the continuous control \mathbf{u} . They fulfil a uniform *Lipschitz condition*¹⁴.

Assumption 2.3: For fixed time t , switching manifolds do not intersect, i.e. $M_{q_i q_j}(t) \cap M_{q_i q_k}(t) = \emptyset$ with $q_j \neq q_k$ for $i, j, k \in Q$.

2.2 Formulation of Hybrid Optimal Control Problems

The saying “A problem well put is a problem half solved” may be an exaggeration, but it shows the importance of problem formulation. Three elements are required to formulate an OCP as well as a HOCP: 1) a mathematical description of the system to be controlled, 2) a set of constraints of the system, 3) a measure to evaluate the performance.

A hybrid system \mathbb{H} is formulated in the previous subsection to provide the mathematical description. At the same time, the collection \mathcal{C} provides the constraints for the HOCP. In what follows, performance measures and boundary conditions are discussed. The HOCPs are formulated.

2.2.1 Performance Measures for Hybrid Optimal Control Problems

The performance measure is selected by the hybrid optimal control designer based on the goal to be reached, e.g. to move an object to a given position with least time. In the case of EMSs, the goal is often to have the least energy consumed while following a given speed profile—the continuous states are normally related to it—over a given time interval. With the dependency on the time in mind, the performance of a hybrid system \mathbb{H} can be measured in the form

¹² A nonempty set E is said to be bounded, if it fulfils The Completeness Axiom [44, p. 9] The definition of Closed Set can be found in [44, pp. 16-17]

¹³ Let X and Y be two measurable sets equipped with respective σ -algebra Σ_1 and Σ_2 . A function $f: X \rightarrow Y$ is said to be measurable, if for all $E \in \Sigma_2$, $f^{-1}(E) := \{x \in X \mid f(x) \in E\} \in \Sigma_1$.

¹⁴ The function f is said to be Lipschitz continuous on a set E , if there is a $L \geq 0$ for which $\|f(x') - f(x)\| \leq L \cdot \|x' - x\|$ for all $x', x \in E$ [44, p. 25].

$$\begin{aligned}
J = m(\mathbf{x}(t_f)) &+ \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t), q(t), \varpi(t)) dt \\
&+ \sum_{j=1}^{N_{\text{exe}}} l_{\text{switch}}(\mathbf{x}(t_j), q(t_j), \varpi(t_j)),
\end{aligned} \tag{2.16}$$

where t_0 and t_f are the initial and final time. m , l and l_{switch} are scalar functions, which are called *endpoint function*, *instantaneous cost* and *switching cost* respectively. The endpoint function assigns a real value to the continuous states at the final time. The instantaneous cost is a function that assigns a real value to a point $(\mathbf{x}(t), \mathbf{u}(t), q(t), \varpi(t))$. The switching cost is a function that assigns a real value to a point $(\mathbf{x}(t_j), q(t_j), \varpi(t_j))$, where t_j is the time instance of a switching, indicating the cost caused by an autonomous or controlled switch. In [45], the endpoint function is also called as an *endpoint functional* or *Mayer functional*. The integral of the instantaneous cost is called an *integral functional*. A functional consisting of both functionals is called a *Bolza functional*.

It is necessary to review briefly the definition of a function, so that the concept of functionals is easier to be understood.

Definition 2.8 *function* [46]: Let $S1$ and $S2$ be two sets. A function f from $S1$ to $S2$, i.e. $f: S1 \rightarrow S2$, is a relation between them that for each $s_1 \in S1$, there is one and only one associated $s_2 \in S2$. The set $S1$ is called the *domain* of the function, which is written as $\mathbf{dom}S1$. The set $S2$ is called its *range*.

△

A scalar function assigns a real number to each point or vector in its domain. The definition of a functional parallels that of a function.

Definition 2.9 *functional* [16]: Let F be a class of functions and S be a set of real number. A functional J from F to S is a relation between them that for each function $f \in F$, there is one and only one associated $s \in S$. The class F is called the *domain* of the functional. The set S is called its *range*.

△

Loosely speaking, one may consider a functional a “scalar function of a function”.

With the idea in mind, that a functional measures function, the arguments of J in (2.16) are complemented with the *Boundary conditions* and the functions defined in **Definition 2.7** as

$$\begin{aligned}
J(\mathbf{x}(t_0), \mathbf{u}(\cdot), q(t_0), \varpi(\cdot)) &= m(\mathbf{x}(t_f)) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t), q(t), \varpi(t)) dt \\
&+ \sum_{j=1}^{N_{\text{exe}}} l_{\text{switch}}(\mathbf{x}(t_j), q(t_j), \varpi(t_j)),
\end{aligned} \tag{2.17}$$

which says that the Bolza functional J assigns a performance measure to the quadruple consisting of the initial continuous states, the initial discrete state, the continuous control function, and the discrete control function. Often, the initial discrete state and the continuous states are omitted in the argument of the functional and supplied in boundary conditions.

In the rest of the work, a functional measuring the performance is called a *cost functional*, regardless of its composition and type.

2.2.2 Boundary Conditions

Boundary conditions define initial and / or final values of the desired trajectories. For a system of ordinary differential equations, it is usually necessary to predefine one boundary condition per differential equation. For a HOCP, up to $2N_x$ number of boundary conditions can be predefined, due to the necessary conditions provided by the Hybrid Minimum Principle described later. Boundary conditions can be defined as

Definition 2.10 *Boundary Conditions*: The boundary conditions $\psi_0(\cdot)$ imposes N_{ψ_0} equality constraints on the initial continuous states $\mathbf{x}(t_0)$, while the boundary conditions $\psi_f(\cdot)$ imposes N_{ψ_f} equality constraints on the final continuous states $\mathbf{x}(t_f)$ as

$$\psi_0(\mathbf{x}(t_0)) = \mathbf{0}, \tag{2.18}$$

$$\psi_f(\mathbf{x}(t_f)) = \mathbf{0}. \tag{2.19}$$

A coupled boundary conditions is defined as

$$\psi(\mathbf{x}(t_0), \mathbf{x}(t_f)) = \mathbf{0}, \tag{2.20}$$

if both the initial and the final continuous states are constrained by equality constraints.

In the context of the EMS problem, the boundary conditions are most often linear and can be stated as

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad (2.21)$$

$$\mathbf{x}(t_f) = \mathbf{x}_f. \quad (2.22)$$

2.2.3 Problem Formulation

For a hybrid optimal control problem, the goal is to find the continuous control function $\mathbf{u}(\cdot) \in \mathcal{U}(\mathbf{x}(\cdot), q(\cdot))$ and the discrete control function $q(\cdot) \in \mathcal{B}_{q(\cdot)}$, such that 1) all constraints from the set \mathcal{C} are met; 2) the boundary conditions are satisfied; 3) the cost functional is minimized.

The optimal control problem for a hybrid system can be stated as follows:

Definition 2.11 *Hybrid Optimal Control Problem* [41, p. 14]: Given a hybrid system \mathbb{H} by **Definition 2.1**, an optimal control problem stated with a cost functional (2.17) as

$$\begin{aligned} J(\mathbf{u}^*(\cdot), \varpi^*(\cdot)) = & \min_{\mathbf{u}(\cdot) \in \mathcal{U}(q(\cdot)), \varpi \in \mathcal{B}_{q(\cdot)}} m(\mathbf{x}(t_f)) \\ & + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t), q(t), \varpi(t)) dt \\ & + \sum_{j=1}^{N_{\text{exe}}} l_{\text{switch}}(\mathbf{x}(t_j), q(t_j), \varpi(t_j)) \end{aligned} \quad (2.23)$$

subject to

$$\dot{\mathbf{x}}(t) = \mathbf{f}_{q(t)}(\mathbf{x}(t), \mathbf{u}(t)), \quad t \in [t_0, t_f] \quad (2.24)$$

$$q(t_j^+) = \Pi(\mathbf{x}, q(t_j^-), \varpi(t_j)), \quad t_j \in t \quad (2.25)$$

$$\mathbf{x}(t_j^+) = \mathbf{x}(t_j^-) + A_{q(t_j^-), q(t_j^+)}(\mathbf{x}(t_j^-)) \quad (2.26)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (2.27)$$

$$\mathbf{x}(t_f) = \mathbf{x}_f \quad (2.28)$$

$$q(\cdot) \in \mathcal{Q}, \mathbf{x}(\cdot) \in \mathcal{X}(q(\cdot)). \quad (2.29)$$

△

For hybrid systems without state-jumps, the HOCP is stated as follows:

Definition 2.12 *Hybrid Optimal Control Problem without State-Jumps* : Given a hybrid system \mathbb{H}_1 by **Definition 2.4**, an optimal control problem stated with a cost functional (2.17) as

$$J(\mathbf{u}^*(\cdot), \varpi^*(\cdot)) = \min_{\mathbf{u}(\cdot) \in \mathcal{U}(q(\cdot)), \varpi \in \Pi_{q(\cdot)}} m(\mathbf{x}(t_f)) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t), q(t), \varpi(t)) dt \quad (2.30)$$

subject to

$$\dot{\mathbf{x}}(t) = \mathbf{f}_{q(t)}(\mathbf{x}(t), \mathbf{u}(t)), \quad t \in [t_0, t_f] \quad (2.31)$$

$$q(t_j^+) = \Pi(\mathbf{x}, q(t_j^-), \varpi(t_j)), \quad t_j \in \mathcal{t} \quad (2.32)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (2.33)$$

$$\mathbf{x}(t_f) = \mathbf{x}_f \quad (2.34)$$

$$q(\cdot) \in \mathcal{Q}, \mathbf{x}(\cdot) \in \mathcal{X}(q(\cdot)). \quad (2.35)$$

The switching cost functions are omitted, due to the absence of state-jumps. The cost caused by switching is considered in the integral functional.

△

This work focuses on the solution and the applications related to HOCs without state-jumps, since it is the class of HOCs relevant to the hybrid system in this work.

For the hybrid optimal control problem with state-jumps, more regularity conditions need to be considered to ensure the existence and uniqueness of the solution [47].

2.3 Discretization and Integration Schemes

If the *ordinary differential equations* (ODE) (2.31) describing the system dynamics cannot be solved analytically, the *initial value problems* (IVP) has to be formed on discrete-time and its numerical solution is fundamental to the solution of HOC.

The time t in the time interval $[t_0, t_f]$ can be discretized into N_t steps, as

$$t_{[k]} = t_0 + k \cdot h_{[k]}, \quad k = 0, 1, 2, \dots, N_t, \quad (2.36)$$

$$h_{[0]} = 0, \quad (2.37)$$

$$h_{[k]} = t_{[k]} - t_{[k-1]}, \quad k = 1, \dots, N_t, \quad (2.38)$$

where k is the time index and h_k is the *step-length*. The squared brackets stress that the integer

in-between indexes a point on a discretization grid, may it be time or continuous states. As a result, a discretization grid \mathcal{G}_t

$$\mathcal{G}_t = \{t_{[0]}, t_{[1]}, t_{[2]}, \dots, t_{[N_t]}\} \quad (2.39)$$

is obtained. A discretization scheme approximates the exact solution $\mathbf{x}(t_{[k]}) = \mathbf{x}_{[k]}$ on the discretization grid. For the ODE (2.31) with given initial condition $\mathbf{x}_{[k]}$, the simplest discretization scheme, namely Euler method, approximates the exact solution $\mathbf{x}(t_{[k+1]})$ by

$$\mathbf{x}(t_{[k+1]}) = \mathbf{x}_{[k]} + \int_{t_{[k]}}^{t_{[k+1]}} \mathbf{f}_{q(t)}(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (2.40)$$

$$\mathbf{x}_{[k+1]} = \mathbf{x}_{[k]} + h_{[k]} \cdot \mathbf{f}_{q_{[k]}}(\mathbf{x}_{[k]}, \mathbf{u}_{[k]}), \quad k = 0, 1, 2, \dots, N_t - 1, \quad (2.41)$$

$$\mathbf{x}_{[0]} = \mathbf{x}(t_{[0]}). \quad (2.42)$$

Runge-Kutta (RK) methods [48] are often used to approximate the integral in (2.40). RK methods with different orders are widely discussed in the context of optimal control problems [49–51]. One property that attracts the attention is the adjoint consistency of RK methods [52]: RK methods provide a discretization grid whose discrete adjoint¹⁵ is a consistent approximation of the infinite¹⁶ dimensional adjoint function. Based on this advantage, RK methods are applied as integration scheme in this work. The fourth order RK integration scheme (RK4) [53, pp. 15-37] follows the following steps:

$$\begin{aligned} \mathbf{x}_{[k+1]} &= \mathbf{x}_{[k]} + h_{[k]} \cdot \Xi_f(\mathbf{x}_{[k]}, q_{[k]}, \mathbf{u}_{[k]}, h_{[k]}), \quad k = 0, 1, 2, \dots, N_t - 1, \\ &= \mathbf{x}_{[k]} + \frac{h_{[k]}}{6} \cdot (\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4), \end{aligned} \quad (2.43)$$

$$\mathbf{k}_1 = \mathbf{f}_{q_{[k]}}(\mathbf{x}_{[k]}, \mathbf{u}_{[k]}), \quad (2.44)$$

$$\mathbf{k}_2 = \mathbf{f}_{q_{[k]}}\left(\mathbf{x}_{[k]} + h \cdot \frac{\mathbf{k}_1}{2}, \mathbf{u}_{[k]}\right), \quad (2.45)$$

$$\mathbf{k}_3 = \mathbf{f}_{q_{[k]}}\left(\mathbf{x}_{[k]} + h \cdot \frac{\mathbf{k}_2}{2}, \mathbf{u}_{[k]}\right), \quad (2.46)$$

$$\mathbf{k}_4 = \mathbf{f}_{q_{[k]}}(\mathbf{x}_{[k]} + h \cdot \mathbf{k}_3, \mathbf{u}_{[k]}), \quad (2.47)$$

¹⁵ An efficient numerical way to compute the gradient for optimal control.

¹⁶ A continuous time interval leads to an infinite dimensional vector of time.

where Ξ_f denotes an operator that uses RK method to approximate the integral of function f in (2.40). \mathbf{k} denotes the estimation of the function f in the interior of the time interval $[t_0, t_f]$. Figure 2.3 illustrates an example of using RK4 method to estimate $x_{[k+1]}$ for a one-dimensional (1-D) dynamic system. k_1 - k_4 are sequentially calculated. $x_{[k+1]}$ from (2.43) for this example is 1.9754, whose relative error is 6.90×10^{-5} , compared to the analytical result.

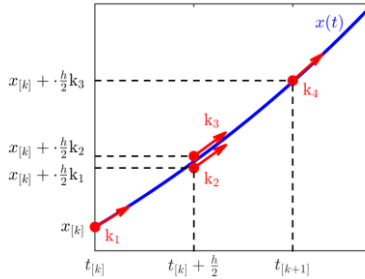


Figure 2.3: RK4 integration scheme performed on a 1-D system, whose dynamic is $\dot{x} = 0.4x + 0.1$. Analytical solution is $x(t) = e^{0.4t+C} - 0.04$ with $C = 0$, $t_{[k]} = 1$ and $h = 0.1$.

2.4 State of the Art Solution Methods

After the hybrid systems defined with necessary prerequisites in Section 2.1 and the HOCP stated in Section 2.2, this section gives an overview of the solution methods, followed by a general comparison of their properties. The methods are categorized in three main types: dynamic programming (DP), direct methods (DM), and indirect methods (IM). They are discussed with a focus on the hybrid system \mathbb{H}_1 by **Definition 2.4**.

2.4.1 Dynamic Programming

The DP was formulated for continuous nonlinear systems by Bellman¹⁷ in 1960s based on Bellman's optimality principle, which states "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision" [54, pp. 81-86]. It is extended to HOCPs with autonomous and controlled switching in [56]. Optimal control law for hybrid systems with controlled switching is approximated based on DP in [57] and [58].

¹⁷ Original work of Bellman see [54]. There are excellent textbooks on DP, for instance, Bertsekas [55] and Kirk [16].

The discrete DP, i.e. the DP with a system operating on discrete time, is often referred to as “DP” in the literature, since it is the natural implementation on a digital computer [16, p. 86]. Its core is called *Bellman equation*, whose variation for hybrid systems can be stated as following recurrence relation

$$V_{[k-1]}(\mathbf{x}_{[k-1]}, q_{[k-1]}) = \min_{\mathbf{u}^{(c)} \in \mathcal{U}(q^{(c)}), \bar{\omega} \in \mathcal{B}_{q^{(c)}}} \{h \cdot l(\mathbf{x}_{[k-1]}, \mathbf{u}_{[k-1]}, q_{[k-1]}, \bar{\omega}_{[k-1]}) + V_{[k]}(\mathbf{x}_{[k]}, q_{[k]})\}, \quad (2.48)$$

$$\mathbf{x}_{[k]} = \mathbf{x}_{[k-1]} + h \cdot \Xi_f(\mathbf{x}_{[k-1]}, q_{[k-1]}, \mathbf{u}_{[k-1]}, h), \quad (2.49)$$

$$q_{[k]} = q_{[k-1]} + \bar{\omega}_{[k-1]}, \quad (2.50)$$

where $V(\mathbf{x}_{[k-1]}, q_{[k-1]})$ denotes the *value function* of the hybrid state $(\mathbf{x}_{[k-1]}, q_{[k-1]})$ at time $t_{[k-1]}$. A value function is the smallest possible value of the performance measure of the subproblem that initiate with $(\mathbf{x}_{[k-1]}, q_{[k-1]})$ starting with time $t_{[k-1]}$. If the performance measure represents a cost that is to be minimized, which is true for this work, a value function is also referred as *cost-to-go* function [59]. Once the value function at each time point is determined, the corresponding controls represent the optimal control.

In order to have (2.48)-(2.50) solved, the continuous states are discretized on the discretization grid $\mathcal{G}_x^{N_x}$. Together with the discrete state and the discretized time, a discretization grid $\mathcal{G}_x^{N_x} \times \mathcal{Q} \times \mathcal{G}_t$ is generated. At the same time, the continuous controls are discretized, and together with the admissible discrete control set, a discretization grid $\mathcal{G}_u^{N_u} \times \Pi_q$ is generated. Figure 2.4 illustrates a discretization grid of a hybrid system, on which the value function is calculated based on the Bellman equation. At each time point, the 1-D continuous states and the discrete state form a discretization grid $\mathcal{G}_x \times \mathcal{Q}$, $\mathcal{Q} := \{q_1, q_2, q_3, q_4\}$, on which the hybrid state is a pair $\{x_{[i]}, q_j\}$ ¹⁸ with i indexing the points on \mathcal{G}_x and j indexing the elements out of \mathcal{Q} . In conjuncture with the discretization of time, a pair $\{x_{[i],[k]}, q_{j,[k]}\}$ indicates any point on the discretization grid of the hybrid system. To have the value function (2.48) of the hybrid state $\{x_{[3]}, q_1\}$ at time $t_{[k-1]}$ determined, the state transitions, i.e. (2.49) and (2.50), from $[k-1]$ to $[k]$ are calculated based on the discretization grid of controls $\mathcal{G}_u^{N_u} \times \Pi_{q_1}$ (not shown in the figure), which are marked as red and blue arrows. At the same time, the cost function l of each state transition is calculated by inserting the hybrid state and the controls into $l(\mathbf{x}_{[3],[k-1]}, \mathbf{u}_{[m],[k-1]}, q_{1,[k-1]}, \bar{\omega}_{n,[k-1]})$, in which m and n index the discretization grid of controls. The value functions on the discretization grid $\mathcal{G}_x \times \mathcal{Q}$ at time $[k]$ are already known from last step of recursion. The Bellman equation to determine the value function

¹⁸ q_j does not contain square brackets, since j indexes the element out of \mathcal{Q} .

at time $[k - 1]$ is completed.

In Figure 2.4, the continuous states at time $[k]$ calculated by (2.49) are marked as crosses, which are not necessarily at the discretization grid $\mathcal{G}_x \times \mathcal{Q}$ at time $[k]$. The value function of them is approximated with *nearest-neighbour* or *interpolation* of the value function on the discretization grid $\mathcal{G}_x \times \mathcal{Q}$ at time $[k]$. The accuracy of the approximation, naturally, is dependent on the fineness of the discretization. The calculated continuous states can fall out of the admissible continuous state space \mathcal{X} . A penalty is in practice assigned to (2.48), which requires delicate attention to ensure accuracy, if the value function has big gradient at the boundary of the admissible space [59].

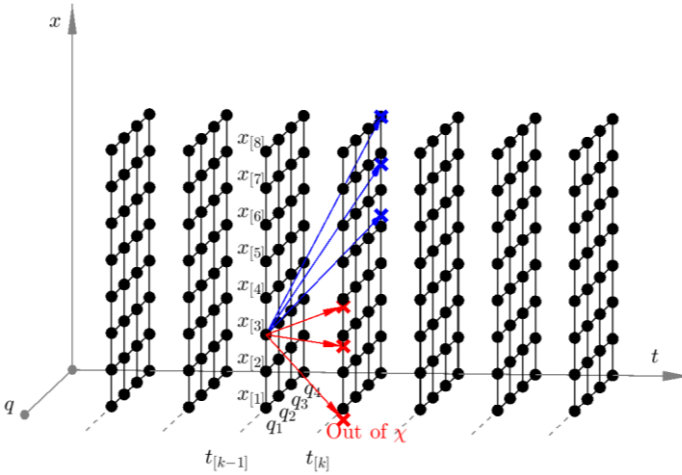


Figure 2.4: An example of discrete Bellman equation on a discretization grid $\mathcal{G}_x \times \mathcal{Q} \times \mathcal{G}_t$. A transition to discrete state q_2 or q_3 is not admissible, which is defined by the admissible discrete control set \mathcal{B}_{q_1} .

The *Continuous DP*, i.e. the DP with a system operating on continuous time, transfers the original optimal control problem to a nonlinear partial differential equation system [16, p. 86]. Since it is less relevant to most modern applications on digital computers, the discussion of it is omitted. See [16, pp. 86-100] for more information.

2.4.2 Indirect Methods

Indirect methods root in *Calculus of Variations*, which traces back to Johann Bernoulli's solution to his challenge to his peers of the Brachystochrone Problem¹⁹ at the end of 17th century (see [61])

¹⁹ To find a curve lying on the plane between a point A and a lower point B that is not directly below A in such

for a survey of the historical perspectives of optimal control). In 1950s, though considered a minor addition to calculus of variations and not highly regarded at the time, the *Pontryagin Minimum Principle*²⁰ was developed in the Soviet Union [62, pp. 9-69]. Conventional wisdom holds it to be the birth of optimal control, since it is the first that addresses constrained states and controls, which are major concerns of control problems. Instead of minimizing a functional that integrates cost function over time, a continuous OCP is solved by finding the continuous control function and the continuous state trajectory that satisfy the optimality conditions. The optimality conditions lead to a Two-Point or Multi-Point Boundary Value Problem (TPBVP or MPBVP)²¹, which includes differential equations of continuous states and *costates*²², algebraic equations of continuous controls, boundary conditions of continuous states, *transversality condition*²³ and time.

Remark: costates are introduced for mathematical purpose, which have no physical meaning [16, pp. 161-166]. In a continuous OCP, a functional of $N_u + N_x$ functions, i.e. $\mathbf{u}(\cdot)$ and $\mathbf{x}(\cdot)$, is to be minimized. However, only $\mathbf{u}(\cdot)$ is independent. Costates are necessary to provide further optimality conditions for indirect methods, so that Boundary Value Problems are solvable. They can be interpreted as Lagrange multiplier that are introduced in optimization problems to include constraints into cost function [63, pp. 215-232]. When the OCP is solved, the corresponding costates have strong correlation with the gradient of the value function in continuous DP [62, pp. 69-75].

The Minimum Principle for hybrid cases were derived by Sussmann for a general class of systems [20]. HOCP of Hybrid systems that share similar definition with **Definition 2.1** were studied in [64, 65]. For hybrid systems with strictly continuous states that are similar to hybrid systems \mathbb{H}_1 defined by **Definition 2.4**, Shaikh [66] and Xu [67] have derived optimality conditions and conceptual algorithms. It was shown in [68] that Hybrid Minimum Principle is an extension of Pontryagin Minimum Principle. More detailed review on HMP can be found in [69, pp. 159-163].

The Theorem of the Hybrid Minimum Principle (HMP) [68]: consider a hybrid system \mathbb{H}_1 by **Definition 2.4** and define the family of indexed Hamiltonians for HOCP (2.30)

a way that a bead slides from A to B under gravity without friction within shortest time [60]. Giants such as Newton, Leibniz, l'Hopital and Jakob Bernoulli participated the challenge and published their solutions.

²⁰ In Pontryagin's original work, the theorem was stated as the Maximum Principle. They are equivalent to the Minimum Principle, since to minimize a function or functional is merely to maximize a function or functional that is formed by putting a negative sign in front of the one that is to be minimized.

²¹ To solve (mostly) first order differential equations with boundary conditions specified at two or more distinct points.

²² Costates are also called adjoint states in some literature [41, p. 32].

²³ Transversality condition is the boudnary condition of states and costates.

$$\mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q) = l_q(\mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}_q^T \cdot \mathbf{f}_q(\mathbf{x}, \mathbf{u}), \quad (2.51)$$

where $\boldsymbol{\lambda}_{q(t)} \in \mathbb{R}^{N_x}$ are indexed costates and T denotes transpose operation. The solution to HOCPC (2.30) satisfies following conditions:

The continuous state trajectory corresponding the optimal control $\mathbf{x}^*(\cdot)$ and the costates $\boldsymbol{\lambda}_{q(t)}^*(\cdot)$ follow

$$\dot{\mathbf{x}}^*(t) = \frac{\partial \mathcal{H}_{q(t)}}{\partial \boldsymbol{\lambda}_{q(t)}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}_{q(t)}(t)) \quad (2.52)$$

$$= \mathbf{f}_{q^*(t)}(\mathbf{x}^*(t), \mathbf{u}^*(t)),$$

$$\dot{\boldsymbol{\lambda}}_{q(t)}^*(t) = -\frac{\partial \mathcal{H}_{q(t)}}{\partial \mathbf{x}_{q(t)}}(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}_{q(t)}^*(t)), \quad (2.53)$$

for $t \in [t_j, t_{j+1}), j = 0, 1, \dots, N_{swt}$ with the boundary conditions $\mathbf{x}^*(t_0) = \mathbf{x}_0$ and $\mathbf{x}^*(t_f) = \mathbf{x}_f$.

The Hamiltonian minimum condition,

$$(\bar{\omega}^*(t), \mathbf{u}^*(t)) = \underset{\bar{\omega}(\cdot) \in \mathcal{B}(q(\cdot)), \mathbf{u} \in \mathcal{U}}{\operatorname{argmin}} \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q), \quad (2.54)$$

holds for $t \in [t_j, t_{j+1}), j = 0, 1, \dots, N_{swt}$.

At the final time t_f , the transversality condition,

$$\boldsymbol{\lambda}_{q(t_f)}(t_f) = \nabla_{\mathbf{x}} m(\mathbf{x}^*(t_f))^T, \quad (2.55)$$

holds.

At the time t_j , the discrete state switches. The following conditions are satisfied

$$\boldsymbol{\lambda}_{q(t_j^-)}(t_j^-) = \boldsymbol{\lambda}_{q(t_j^+)}(t_j^+), \quad (2.56)$$

$$\mathcal{H}_{q(t_j^-)}(\mathbf{x}(t_j^-), \mathbf{u}(t_j^-), \boldsymbol{\lambda}_{q(t_j^-)}(t_j^-)) = \mathcal{H}_{q(t_j^+)}(\mathbf{x}(t_j^+), \mathbf{u}(t_j^+), \boldsymbol{\lambda}_{q(t_j^+)}(t_j^+)). \quad (2.57)$$

△

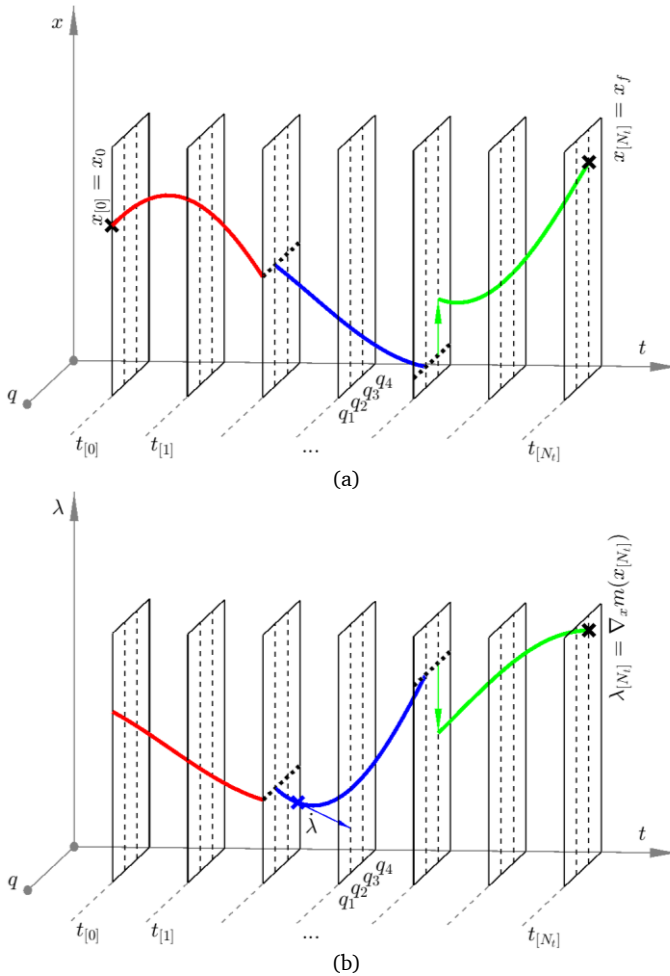


Figure 2.5: An example of a solution with HMP to a HOCP with a 1-D continuous state. (a) shows continuous state trajectory. (b) shows costate trajectory.

Figure 2.5 shows an example of 1-D continuous state and costate trajectories of a solution determined by HMP on a discretized time grid \mathcal{G}_t with N_t steps. 1-D continuous control is not shown in the figure. Unlike discretize DP, continuous states and continuous controls are not discretized. The time derivatives \dot{x} and $\dot{\lambda}$ at any time point are determined by differential equations (2.52) and (2.53), in which ω^* and u^* are obtained by minimizing the Hamiltonian \mathcal{H}_q at the time point (2.54). By doing so, both trajectories evolve from $t_{[0]}$ to $t_{[N_t]}$. As long as $\omega_{[k]}^*$ is not zero, q switches

to a new value determined by $\bar{\omega}_{[k]}^*$, as the system changes from q_1 (red curve) to q_2 (blue curve) and q_2 to q_3 (green curve) at different time points in Figure 2.5. (2.52) and (2.53) are indexed, which says that the dynamics of the system depend on q . State-jumps happen, when q switches from q_2 to q_3 , which exemplifies the situation that state-jump functions defined in a hybrid system are not empty. As long as a solution is obtained, the boundary conditions of x at $t_{[0]}$ and $t_{[N_t]}$ are fulfilled and λ at $t_{[N_t]}$ finds the transversality condition²⁴ defined by (2.55). Importantly, $\lambda_{[0]}$ is unknown. Therefore, to obtain a solution with HMP is to solve a TPBVP of $\lambda(\cdot)$ and $x(\cdot)$, whose time derivatives follow (2.52) and (2.53) with boundary conditions of

$$x(t_0) = x_{[0]} = x_0, \quad (2.58)$$

$$x(t_f) = x_{[N_t]} = x_f, \quad (2.59)$$

$$\lambda(t_f) = \lambda_{[N_t]} = \frac{\partial m(x(t_f))}{\partial x} = \frac{\partial m(x_{[N_t]})}{\partial x}. \quad (2.60)$$

The *Shooting method* is often applied to solve a BVP [69, p. 218]. In principle, it proposes iterative initial values, starting with which the time derivatives are integrated from the initial time point to the final time point, until the value at the final time point converges to the terminal condition. Formally, a guess $\hat{\lambda}$ is proposed to form the initial value of the extended state vector $y := (x^T, \lambda^T)^T$, i.e.

$$y(t_0) = \begin{pmatrix} x_0 \\ \hat{\lambda} \end{pmatrix}, \quad (2.61)$$

in such a way that the nonlinear equations²⁵,

$$y(t_f) - \begin{pmatrix} x_f \\ \frac{\partial m(x(t_f))}{\partial x} \end{pmatrix} = \mathbf{0}, \quad (2.62)$$

are solved. Figure 2.6 schematically exemplifies the shooting method, which acquires different $x(t_f)$ and $\lambda(t_f)$ with different $\hat{\lambda}$. After four iterations modifying $\hat{\lambda}$, (2.62) is solved. To have the illustration easy to comprehend, discrete state remains constant and is omitted in Figure 2.6. In

²⁴ Should a cost functional not include an endpoint function m , λ would have to end with 0.

²⁵ $\mathbf{0}$ in the equation is a $2N_x$ -dimensional zero vector.

most cases, sequence of discrete state changes at different iterations, which further complicates the process to solve (2.62), as the function of time derivatives (2.52) and (2.53) changes, due to the change of discrete state.

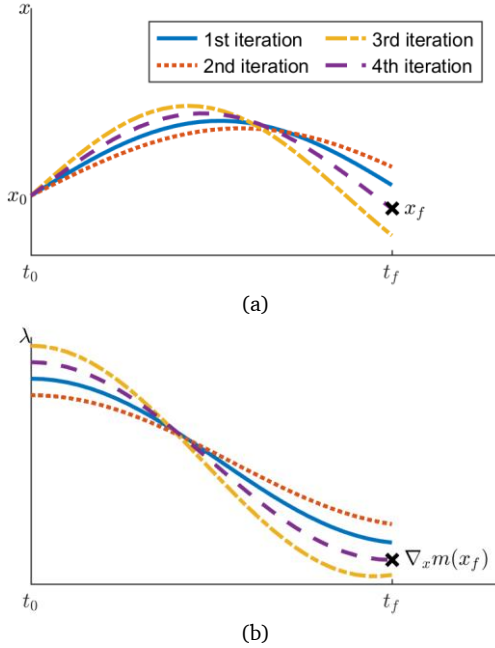


Figure 2.6: Schematic of shooting method to solve a TPBVP transformed from a HOCP of 1-D continuous state by HMP. (a) continuous state trajectories in different iterations. (b) costate trajectories in different iterations.

In HMP, the minimization (2.54) is often not solved analytically. A numerical minimization of the instantaneous Hamiltonian is necessary to determine u_q^* with appropriate methods, e.g. *Sequential Quadratic Programming* (SQP)²⁶, after which ϖ^* can be determined [69, p. 229]. The numerical minimization is non-trivial for the solution of the HOCP, since 1) their precisions are strongly connected²⁷, and 2) their computational complexities are strongly connected²⁸. However,

²⁶ SQP is a class of algorithms for solving nonlinear programming (NLP) problems by iteratively solving a sequence of optimization subproblems. See [70, pp. 529-562] for more information of it.

²⁷ If \mathcal{H} has multiple local minima, the minimization can be hard to converge.

²⁸ [69, p. 229] suggests evaluating \mathcal{H} on a grid and using the smallest value to initiate SQP, so that the precision of the minimization can be improved. This, on the other hand, increases the computational complexity.

the minimization is rarely investigated, either in theoretical studies—which is understandable, since they focus on optimal control but not optimization—or in application studies.

Roots of (2.62) can be particularly difficult to find, since costates offer generally no physical interpretation, which makes even the order of magnitude of $\hat{\lambda}$ hard to be approximated. For the most simple cases, where 1-D state or 1-D costate can be removed from (2.62), *regula falsi* methods, e.g. *bisection*, *scent* and *Pegasus* [71], are recommended for the resulting 1-D TPBVP [69, p. 230]. For (2.62) with higher dimension, *Newton type* methods are classical methods to consider [69, p. 219]. *Quasi-Newton* methods avoid computations of Jacobian and Hessian, which makes it generally operates with higher efficiency than Newton type methods for root finding problems [72, pp. 49-56]. Similar to the minimization of *instantaneous Hamiltonian*, few works mentioned the method applied to find the root of (2.62): Passenberg applied Newton method in [41, 72; 79]; Boehme applied regula falsi methods in [73]. Furthermore, the initialization of $\hat{\lambda}$ is non-trivial but lacks investigation. Passenberg proposed two initialization concepts: 1) use the results of Direct Method to generate $\hat{\lambda}$ [41, pp. 80-81]; 2) use the results of *Min-H* method, a gradient method that finds local minima, to generate $\hat{\lambda}$ in [41, pp. 87-110] and [74]. Both concepts have HOCP solved with optimization methods to ensure the solution of (2.62). A simpler initialization method is necessary, if the efficiency of the solution with HMP is desired, especially for an online EMS.

2.4.3 Direct Methods

With direct methods, the HOCP is treated as a *Mixed-Integer Nonlinear Programming* (MINLP)²⁹ problem, which optimizes the cost functional directly, as the name of the methods indicate. The continuous parts in the original HOCP, i.e. time, continuous controls and continuous states³⁰, are discretized. The problem is transformed to a finite-dimensional MINLP.

Figure 2.7 exemplifies the discretization through single direct shooting³¹ of a hybrid system defined by **Definition 2.4** with 1-D continuous state and 1-D continuous control on the time grid \mathcal{G}_t . It is assumed that the discrete control directly changes discrete state, as

$$q(t^+) = \omega(t). \quad (2.63)$$

²⁹ It is obviously a Mixed-Integer Linear Programming, if its prerequisites are fulfilled. However, what concerns this work is nonlinear cost function. Only MINLP is discussed.

³⁰ If shooting methods are used, only continuous controls are discretized. If collocation is used, both continuous states and continuous controls are discretized [75, pp. 123-134].

³¹ See [18, pp. 36-41] for general formulation of direct methods: single direct shooting, multiple direct shooting and collocation. Details regarding constraints are omitted in this example to simplify the formulation, so that a stronger intuitive idea of direct methods can be achieved.

By doing so, discrete control is omitted in the optimization.

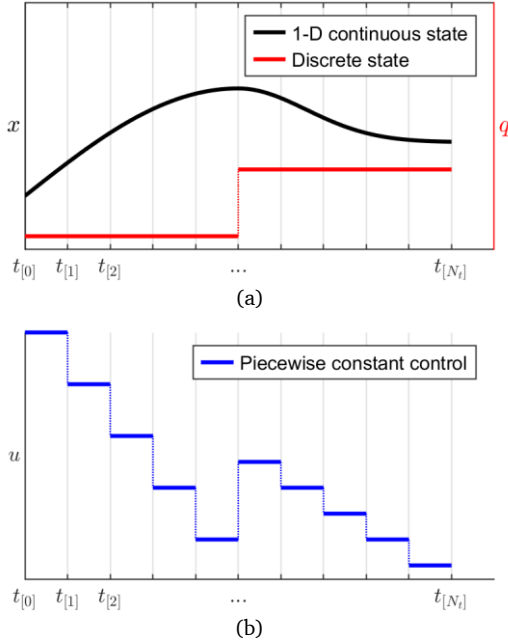


Figure 2.7: Schematic of direct shooting method to a HOCP with 1-D continuous state to MIP. (a) 1-D continuous state and discrete state. (b) Piecewise constant control out of discretized continuous control grid \mathcal{G}_u .

The continuous control space is discretized to form the grid \mathcal{G}_u , and the continuous control function $u(\cdot)$ is approximated with piecewise constant control $\hat{u}(\cdot)$, whose value is selected out of \mathcal{G}_u , i.e.,

$$\hat{u}(t) = \mathcal{G}_u\{p_{[j]}\}, \quad t \in [t_{[j]}, t_{[j+1]}), \quad j = 0, 1, \dots, N_t - 1, \quad (2.64)$$

$$\mathcal{G}_u\{p_{[j]}\} \in \mathcal{G}_u \quad p_{[j]} = 1, 2, \dots, N_{\mathcal{G}_u}, \quad (2.65)$$

where $p_{[j]}$ indexes the point on \mathcal{G}_u , and $N_{\mathcal{G}_u}$ denotes the amount of element in \mathcal{G}_u . The vector $P := (p_{[0]}, p_{[1]}, \dots, p_{[N_t-1]})^T$ encodes $\hat{u}(\cdot)$. The vector $X := (x_{[0]}, x_{[1]}, \dots, x_{[N_t]})^T$ represents the continuous state on \mathcal{G}_t , in which each value is obtained through numerical integration, e.g. RK4 in subsection

2.3. A vector of integer represents discrete state $Q = (q_{[0]}, q_{[1]}, \dots, q_{[N_t]})^T$. The HOCP (2.30) is transformed to the following problem³²

$$\min_{P, Q} \Phi(X, P, Q), \quad (2.66)$$

subject to $G(X, P, Q) = \mathbf{0}, \quad (2.67)$

$$H(X, P, Q) \leq \mathbf{0}. \quad (2.68)$$

The objective function $\Phi(X, P, Q)$ is calculated based on the vectors on the discretization grid and represents the cost functional (2.17). The equality and inequality constraints $G(\cdot)$ and $H(\cdot)$ are the transformations of the constraints in the original HOCP (2.31)-(2.35).

To have the MINLP problem (2.66)-(2.68) solved, it often involves fixing the sequence Q and solving the remaining problem, which has an obvious drawback, as the sequence is typically unknown in advance. A naive approach is to enumerate all possible combination of Q [78, pp. 31-60]. To reduce the enumeration domain, tree search methods such as Branch-and-Bound [79] are applied. See [80] for more advanced methods for same purpose³³. See [18, pp. 86-124] for applications of MINLP in HOCP.

In practice, one searches for available and proper solvers and formulates a MINLP based on the HOCP and the requirements of the solver. See [76] for a review of solvers for convex MINLP.

2.4.4 Comparison among Dynamic Programming, Indirect Method and Direct Method

The advantages and disadvantages of DP, IM and DM for HOCP are discussed. The comparison is conducted regarding four questions: 1) How long does it take to find the solution? 2) What is the property of the solution in theory and how accurate is the solution in practice? 3) How much mathematical knowledge is required? 4) How difficult is the convergence and the initialization? The following outlines the answers to these questions based on [41, pp. 26-30] and [18, pp. 23-50]. Please be aware that this subsection is only a generalized comparison, which might not reflect all properties of a specific method.

³² Formulation of the MINLP transformed from a HOCP is not unified in the literature [18, pp. 85-107, 69, pp. 250-257, 76, 77]. The formulation (2.66)-(2.68) is not necessarily most suitable for problem solving but for conveying a clearer idea of direct methods.

³³ The methods in [80] work on small discretization grids.

- Time complexity³⁴

The time complexity is the total amount of steps taken by an algorithm to execute, as a function of the length of input data [81, pp. 31-67]. As shown in (2.48), the time complexity of DP increases only linearly with the length of time horizon, i.e. N_t in (2.39), and linearly with the number of discrete states with autonomous and controlled switchings N_Q , i.e. the size of the discrete set Q in **Definition 2.1** and **Definition 2.4**. The complexity is not directly correlated to the number of switching N_{exe} in the time horizon (see Section 2.1.1). To have higher accuracy achieved, continuous state space and continuous control space can be discretized finer, which results in larger N_{g_x} and N_{g_u} . DP has a polynomial complexity in N_{g_x} and N_{g_u} . Famously, DP has an exponential complexity in the dimension of continuous state and continuous control, i.e. N_x and N_u , which is called “curse of dimensionality”.

In contrast, DM has a time complexity growing exponentially with N_t for controlled switching, since every discrete state at a time point has to be evaluated. An exception is the algorithm for binary control problem with convexification in [18, pp. 67-84], which has a polynomial complexity. The complexity of DM increases in the form of polynomials with N_Q . DM has a combinatorial complexity in the number of autonomous switchings in the time horizon, while its complexity is not directly correlated to the number of controlled switchings. With increasing accuracy, the time complexity grows with N_{g_x} and N_{g_u} with low polynomial order. It only grows with N_x and N_u in the form of polynomials.

A longer time horizon increases the time complexity of IM in the form of moderate polynomials. The complexity grows linearly with the number of discrete states with controlled switchings but is not directly correlated to the number of controlled switchings. It grows in the form of polynomials with the number of discrete states with autonomous switchings, while IM has a combinatorial complexity in the number of autonomous switchings. The time complexity grows with N_x and N_u in the form of polynomials.

In summary, IM has the lowest time complexity among the three.

- Optimality property in theory and accuracy in practice

A solution found by DP is global optimal, since the Bellman’s optimality and Hamiltonian Jacobian Bellman equation provide sufficient optimality conditions, which is a major strength of DP. The accuracy of the solution of DP is strongly dependent on the fineness of the grid. Due to sampling, its accuracy is relatively low compared to DM and IM, in the case of small

³⁴ The worst case of time complexity is discussed, i.e. big-O notation [81, pp. 31-67]. Linear complexity implies that time complexity increases linearly with the variable n , if $\text{time}(n) = c \cdot n$. Polynomial complexity: $\text{time}(n) = n^c$. Exponential complexity $\text{time}(n) = c^n$. Combinatorial complexity: $\text{time}(n) = c \cdot n!$.

N_{G_x} and N_{G_u} [21]. Additionally, the time complexity of DP grows rather fast, when higher accuracy is desired.

A solution found by DMs is local optimal. DM often provides solution with acceptable accuracy.

A solution found by IMs is local optimal. A major advantage of IMs is its high accuracy. It was reported in [21] that an EMS with IMs achieved same, if not lower, energy consumption compared to an EMS with DP, even when its continuous state space and continuous control space are fine discretized.

- Convergence and initialization

DP algorithms are globally convergent on the discretization grid with an easy initialization with zeros.

DMs have larger convergent domains, compared to IMs. Their initialization is also easier than IMs.

IMs have the smallest convergent domains. Due to the non-intuitiveness of costates [16, pp. 161-166], the initialization of IMs is difficult. Difficulty of convergence and initialization is a major drawback of IMs.

- Required mathematical knowledge

Applying DP requires a lot of mathematical knowledge of optimal control theory.

It is a major advantage of DMs that their applications require the least mathematical knowledge, since their applications mostly focus on finding an appropriate solver and direct transcription, i.e. to transform a HOCp to a MINLP, as it is exemplified in Section 2.4.3.

Applying IMs requires expertise knowledge in optimal control theory and also numerical methods to develop algorithms that find the solution stably. This is a disadvantage of IMs.

2.5 Summary

In this chapter, hybrid systems and HOCps are introduced. The solution methods for HOCps, namely DP, IM and DM, are reviewed. Their applications to HOCps without state-jumps are discussed at length with the goal of helping the reader to develop an idea of the process of their solutions in practice. The solution methods are compared among one another regarding four aspects that are relevant to an online EMS: 1) time complexity, 2) accuracy, 3) convergence and initialization and 4) required mathematical knowledge. Based on their properties and the review of EMS in the next chapter, a solution method will be chosen for the online EMS.

3 State of the Art

As mentioned in Section 1.1, an EMS provides the set-points for the controllers of the power sources³⁵ and the transmission(s) in an electric vehicle in such a way that achieves the minimal energy consumption and some other performance criteria of the vehicle. Often, it operates with a low frequency as low as 1 Hz [82], and considers only longitudinal dynamics [83–85].

Two-drive multi-speed powertrains and their EMSs are relative recent. It is necessary to broaden the horizon of the literature research. EMSs of HEVs and PHEVs are of the most interest for inspiration, since they operate with at least two power sources and are often equipped with multi-speed transmissions. This chapter discusses EMSs of BEVs, HEVs and PHEVs with focuses on: 1) how the continuous and the discrete dynamics were addressed, 2) how the EMS were designed, 3) how are they related to online EMSs for two-drive multi-speed powertrains. The discussion determines the method to be adopted in the online EMS in this thesis. Since a second aspect is to be considered in this work, multi-criteria operation strategies (MCOS) are reviewed. The terminology is introduced to differentiate MCOS and EMS.

Vehicles and Buses are not differentiated. If not explicitly mentioned, the referenced works in this chapter did not consider shift processes. The powertrains mentioned in this chapter include at least one EM. To design an EMS that minimizes energy consumption or fuel consumption is referred to as *the problem*.

3.1 Topology and Optimal Control Problem

Before diving into the review of EMSs, it is necessary to shortly discuss topologies of powertrains from an EMS perspective, since they determine DoFs of powertrain power flows and, therefore, number of controls of EMSs. Furthermore, the differences of OCPs—see Section 2.2.3 for an example—for BEVs, HEVs and PHEVs need to be pointed out so that it can be clear, how EMSs of HEVs and PHEVs discussed in the later subsections are connected to those of BEVs, especially two-drive multi-speed BEVs.

3.1.1 Topology

Typical powertrain topologies of BEVs are presented in Figure 3.1 [10, 86], where the transmissions and the sub-transmissions (ST) can be multi-speed transmissions. A BEV with in-wheel motors requires an EMS with 2 controls, i.e. the torques in the front-wheel EMs and those in the rear-wheel EMs. The OCP is continuous. A BEV with M-drive and N multi-speed transmissions requires

³⁵ Machines that convert other types of energy to mechanical energy, e.g. engines and electric motors.

an EMS with $M + N$ controls. *The problem* contains a discrete part, i.e. gear selection problem, and is a HOCP, if $N > 0$ ³⁶.

See [10, 86] for the analysis of the topologies and a list of series production BEVs with above mentioned topologies. With extra EMs and STs parallelly added into the system, two-drive powertrains can be further extended to multi-drive powertrains.

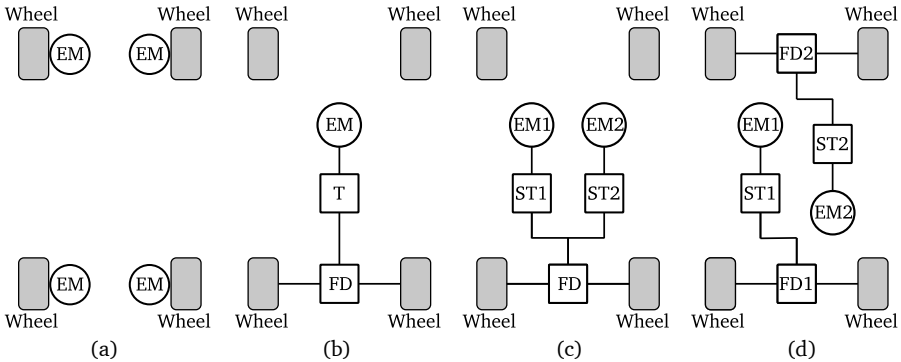


Figure 3.1. Typical powertrain topologies of BEVs. (a): BEV with in-wheel motors. (b) Single-drive BEV. (c): Centralized two-drive BEV. (d): Separated two-drive BEV T: Transmission. ST: Sub-transmission. FD: Final Drive.

There are three main topologies for HEVs and PHEVs: series hybrid, parallel hybrid and series-parallel³⁷ hybrid [83, 88]. Detailed explanation of these topologies can be found in the just-mentioned references. The examples of the vehicles in series production are listed in [83] and [89]. In the series topology, an EM propels the vehicle alone. An engine is used to charge the battery through a generator. *The problem* is continuous, if the EM is connected to a single-speed transmission [24]. In the parallel topology, an EM and an engine can propel the vehicle together³⁸. A multi-speed transmission is necessary for the engine to achieve higher fuel efficiency. A parallel topology as shown in [91] requires two continuous controls and one discrete control. *The problem* is a HOCP. In the series-parallel topology³⁹, an engine and two EMs are connected to the sun gear, the carrier and the ring gear in a planetary gearbox, whose schematic is shown in Figure 3.2, according to the

³⁶ Continuous variable transmission is not considered, due to its disadvantage in efficiency [87].

³⁷ Also called power-slit hybrid.

³⁸ Depending on different relative positions of the EM in the powertrain, there are different configurations under the parallel topology. See [90] for detailed description and analysis.

³⁹ Depending on how the engine and the EMs are configured w.r.t. the planetary gearbox(es) and the output shaft, there are different configurations under the series-parallel topology. See [83] for detailed description.

specific powertrain design [92]. Thanks to the speed coupling mechanism of the planetary gearbox, the series-parallel topology allows the engine speed being varied in favour of higher fuel efficiency without a multi-speed transmission [93]. A series-parallel topology as shown in [94] requires three controls in the EMS and *the problem* is continuous with the preliminaries: 1) The engine speed is regulated instantly and the transient characteristics of the engine is ignored; 2) The planetary gearbox is controlled through its brakes instantly; 3) The rotational inertia of the powertrain does not change. Other works such as [94–96] shared the preliminaries.

Remark: The controls mentioned in the above paragraphs are those that are necessary to determine the powertrain power flows. During the formulation of an OCP, the equality constraints w.r.t the total output torque and the vehicle speed bring certain correlations among the controls, because of which the number of controls in the literature can be one less than described in this subsection.

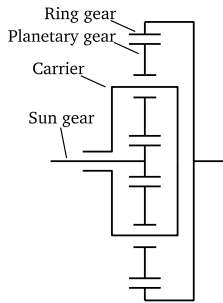


Figure 3.2: Schematic of a planetary gearbox

3.1.2 Optimal Control Problem

In EMSs of HEVs and PHEVs, minimizing fuel consumption is the primary goal. The SOC is considered a state. Its boundary conditions⁴⁰ are to be fulfilled. Rigidly speaking, the vehicle speed should also be part of the states so that the status of a vehicle can be determined. If shift processes are ignored and acceleration does not interrupt, the vehicle speed within a time step can be seen as predefined and removed from the states. An EMS assigns powers for the EM(s) as well as the engine. It also involves gear positions, if a multi-speed transmission exists.

For BEVs, the SOC does not have to be included in the EMS, if its goal is to minimize battery depletion. Minimizing battery depletion is equivalent to minimizing electric energy consumption and, therefore, only electric power needs to be evaluated. For an EMS of a two-drive multi-speed

⁴⁰ For instance, under the charge-sustaining mode, SOC may fluctuate but ought to maintain a certain level over a period of drive. In the most research on HEV EMS, the terminal SOC equals the initial SOC.

BEV that considers shift processes, the angular velocities of both EMs are necessary in the continuous states, while the gear positions are necessary in the discrete states, so that the status of the vehicle can be determined. The boundary conditions of the angular velocities are to be fulfilled. The EMS assigns torques for both EMs and decides the gear position to propel the vehicle following the speed profile. Due to the potential acceleration changes during the shift processes, the vehicle speed within a time step is not predefined.

3.2 Energy Management Strategies

According to review articles, such as [83–85], EMSs can be divided into two basic categories: rule-based EMSs and optimization-based EMSs. Rule-based EMSs are identified as those EMSs that operate based on rules and criteria that has been defined by the system, so that certain performance goals can be achieved. The rules were obtained through heuristics and/or mathematical models. Rule-based EMSs often cost higher energy consumption, compared to optimization-based EMS, for lack of optimization over the whole system considering driving conditions. As the primary goal of the EMS in this thesis is to minimize energy consumption, rule-based EMSs are ruled out from further discussion.

Optimization-based EMSs are those that generate controls through minimizing a cost functional or a cost function. The optimization-based EMSs that require the information of driving cycles as *a priori* are referred to as offline EMSs, while those that do not are referred to as online EMSs⁴¹. Importantly, they are not isolated from each other. The solution methods that have been discussed in length in Section 2.4 all require the information of driving cycles to complete the formulation of the HOCP. The EMS that apply them directly are offline. On the other hand, a group of online EMSs are developed by embedding these solution methods in a Model Predictive Control (MPC) framework, which will be elaborated in Section 3.2.2.1. Another group of online EMSs, Equivalent Consumption Minimization Strategy (ECMS) and its alike, are developed by the simplification of IMs, which will be discussed in Section 3.2.2.2.

Stochastic optimization-based EMSs involve stochastic optimal control. More specifically, in [97–101], stochastic dynamic programming (SDP)⁴² was applied for HEVs and PHEVs. They require too high time complexity (see [103]) to be considered for online EMSs in this thesis.

⁴¹ Online and offline EMSs are also referred to as non-causal and causal EMSs [83].

⁴² Speed profiles and corresponding variables, such as requested power, are modelled as Markov process, whose decision problem is solved by Bellman equation. The expected discounted cost functional in infinite time horizon is minimized [102, pp. 125-132]. It is called the expected total discounted reward in the literature on SDP and markov decision problems.

EMSs for HEVs and PHEVs will be reviewed non-exhaustively under different categories of optimization-based EMS to provide a general overview of the EMS research and inspirations for EMSs for BEVs. After that, EMSs for BEVs will be reviewed.

3.2.1 Offline Energy Management Strategies

3.2.1.1 Dynamic Programming

Given a speed profile, the DP can find the global optimal control functions for an EMS. The discrete DP was used in [24], one of the earliest studies, to tackle *the problem* of a series HEV. To have the boundary condition of the state satisfied, a penalty term that requires parameterization was introduced. The Hamiltonian-Jacobi-Bellman equation⁴³ was applied in [91] to develop an EMS for a two-speed parallel HEV. It provided the first order and the second order optimality conditions to reduce the complexity of calculus. *The problem* included gear selection but ignored shift processes. The boundary conditions of the SOC was incorporated in the cost functional with the help of a Lagrangian multiplier.

The DP was adopted in [104] for a series-parallel PHEV that was modelled as a hybrid system, in which different operating modes were considered discrete states: 1) pure electric propelling, 2) pure electric regenerating, 3) hybrid propelling and 4) electric propelling with battery charging. Worth mentioning, the dynamics in all discrete states were the same and the switching was instant, i.e. shift processes were ignored.

Remark: Though the studies on DP-based EMSs for HEVs and PHEVs have not fully incorporated the goals of this thesis reasoned in Section 1.2, the DP can be used in the EMS in this thesis, as long as the hybrid system can be modelled to meet the goals of this thesis.

3.2.1.2 Indirect Method

Both the DP and the PMP were applied in a comparative study [21] for a parallel HEV with an AMT. The gear selection problem was considered, while the shift processes ignored. The fuel consumption of the HEV controlled by both EMSs were close to each other, while the computation duration of the PMP was less than one fourth of that of the DP. How the TPBVP and the initialization problem were solved was not disclosed. Worth mentioning, the gear position and the shift command, though being discrete, were included in the states and the controls in the formulation of the PMP—a solution method for continuous OCPs. Though not rigorous, the study could still be valid under the assumption that shift processes are neglectable. After all, the HMP is an extension of the PMP (see Section 2.4.2). For a series PHEV in [35], the DP and the PMP were applied to

⁴³ Continuous counterpart of Bellman equation. See [16, pp. 86-90] for details.

generate two offline EMSs. The one based on the PMP outperformed the one based on the DP⁴⁴ w.r.t. the cost evaluated in driving cycle simulations by 0.5 % and the computation duration by several folds.

Under the simplification same as the one for the ECMS (see Section 3.2.2.2), the HMP was adopted in [73] for a parallel HEV that is modelled as a hybrid system. The algorithm was evaluated on several driving cycles, which provided evidence to show that the time complexity of IMs is far less than DMs. The operating modes and the gear positions formed the discrete state. It is noteworthy that, with the simplification, the TPBVP was reduced to 1D. Shift processes were not considered. The initialization of the problem was not disclosed.

Remark: The above-mentioned references provide the evidence for the advantages of IMs mentioned in Section 2.4.4. Applying IMs, more specifically the HMP, for an EMS of a two-drive multi-speed BEV still faces following challenges: 1) A mathematical model including the shift processes needs to be solvable for the HMP; 2) A multidimensional TPBVP needs to be solved, whose solution and initialization in the context of EMSs were not discussed in the literature.

3.2.1.3 Direct method

The EMS of a series HEV was studied in [105] systematically. Under several simplifications⁴⁵, the convex optimization problem in continuous time was formulated, which was further approximated to be a Linear Programming (LP) problem in discrete time. Its equivalent standard LP problem was identified and solved by “PCx”, a readily available software. The offline EMS was used to provide the minimum fuel consumption given the specification of a vehicle and a speed profile.

Designing EMSs for parallel HEVs and PHEVs requires solving HOCPs. One way of simplification to apply NLP is to use a rule-based strategy to deal with the gear selection problem, so that it is separated from the HOCP, for instance in [106] and [107]. The former ignored the engine on/off scenario and applied convex optimization to solve the residual continuous OCP, while the latter considered the engine on/off scenario and applied LP sequentially and iteratively to solve the residual HOCP.

Approaching the HOCPs in the EMSs without separating the gear selection problem was reported. [108–110] proposed to reform the HOCPs to large scale NLP that can be solved by readily available solvers. [77] proposed to solve the HOCPs with certain decomposition techniques, so that readily available solvers for Mixed Integer Linear Programming can be applied. These studies have

⁴⁴ The fineness of the grid was not disclosed.

⁴⁵ 1) The voltage on the electric bus remains constant; 2) The engine transient performances ignored; 3) The battery storage efficiency is constant.

pointed out that the challenging aspect of applying DMs is the high time complexity of solving MINLP.

Remark: Additional to the drawback in time complexity, DMs require readily available solvers that are applicable both in a simulation environment and in a control unit, if an EMS is eventually implemented in a hardware. The latter implementation is nontrivial in practice.

3.2.2 Online Energy Management Strategies

3.2.2.1 Model Predictive Control

An online EMS does not consider speed profiles *a priori*. One natural way to develop it is to use the methods applied in the offline EMSs to solve the OCP on a predicted speed profile in the future based on the current and the history information. It is, more concisely speaking, to embed a solution method in an MPC framework. The EMSs with such a concept are called predictive EMSs in this thesis. Please refer to the book [111] for the theory and the studies of MPC itself.

Figure 3.3 shows the general idea of a predictive EMS. At the current time point j , the driver's request in the prediction horizon, i.e. $j \sim j + N_p$, is predicted based on the current and the past information, so that a predicted trajectory (the red curve with circles) is generated. A solution method is applied to solve the OCP with the predicted trajectory, which results in the control function in the predicted horizon (the light blue stairs). The control function in the control horizon, i.e. $j \sim j + N_c$, is provided by the EMS to the powertrain.

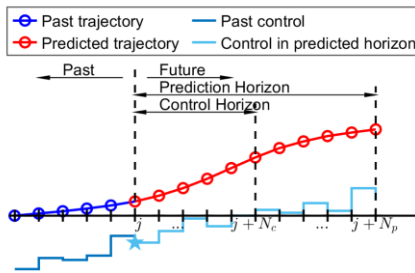


Figure 3.3: Schematic of MPC

Figure 3.4 illustrates the typical components of a predictive EMS for continuous OCPs. The driver's request (r), may it be acceleration, power or torques, in the prediction horizon is predicted in the predictor based on the current driver's request and the outputs of the powertrain (y). The model calculates the outputs in the predicted horizon and the solver determines the optimal control (u^*). Note that the controls at $j + N_c - 1$ forms the states at $j + N_c$. Often, the control horizon is one, which can be seen in the references mentioned later.

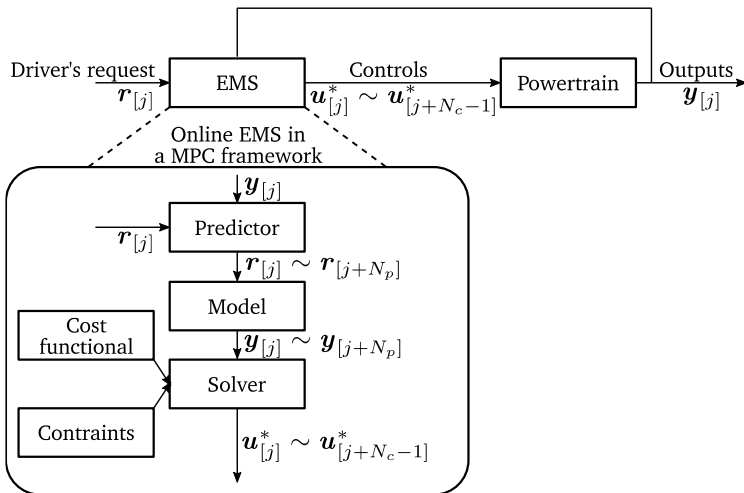


Figure 3.4. Typical components of a predictive EMS for continuous OCPs.

For a series PHEV, a predictive EMS embedded with the PMP was reported in [35]. The pure continuous OCP in the prediction horizon was transformed into a 1-D TPBVP that was solved with the secant method. The driving cycle simulation showed that the predictive EMS with the PMP reduces considerable computation duration but caused 1 % higher cost, compared to the offline EMS based on the DP.

For a parallel HEV, a predictive EMS with a direct method was reported in [112]. The gear selection was separated, as mentioned in Section 3.2.1.3, so that the OCP is continuous. SQP from the optimization toolbox in Matlab/Simulink was used to perform the optimization. Interestingly, it considered the transient characteristics of the engine in the form of look-up tables.

To develop a predictor is nontrivial. Several predictive EMSs with different predictors were developed for a series-parallel HEV and their simulation results were compared against each other in [113]. It was reported that the one with the least prediction error scored the least fuel consumption. The conclusion indicates the importance of a high-performance predictor.

3.2.2.2 Equivalent consumption minimization strategy

The ECMS was originally reported in [22] as an EMS for a parallel HEV. Its relationship to the PMP was analysed in [23] and [114]. By assuming that the dynamics of the SOC is not dependent on the SOC itself, the PMP is simplified and an EMS that only requires the optimization of an instantaneous cost function consisting of the fuel consumption rate and the battery electric power is generated. The equivalent factor in the cost function needs to be parameterized. Adaptive ECMSs

were proposed in [115–117]. They used different variables, such as the SOC and the predictive driving condition, to modify the equivalent factor in a closed-loop.

Remark: Since only the electric energy in the battery concerns an EMS for BEVs, the ECMSs is not in particular interest of this thesis.

3.2.3 Energy management strategies for BEVs

For single-drive multi-speed BEVs, simple EMSs are sufficient, if shift processes are neglected. In [118] and [119], rule-based shift strategies were developed by minimizing instantaneous energy consumption. Offsets based on requested torques were introduced to avoid frequent shifting. For a single-drive two-speed BEV, a predictive EMS with enumeration was developed In [120].

For a two-drive multi-speed BEV, an EMS that minimizes the instantaneous energy consumption was developed in [121]. Gear positions and torques are variables. Under the assumption that shift processes are ignored, the EMS scored indeed the minimum energy consumption. Realistically, the energy consumption might not be optimal, if the EMS is evaluated with a vehicle model that considers shift processes.

An EMS that considers shift processes was developed in [26] for a single-drive two-speed BEV with an automated transmission with planetary gearsets developed in [122] and [123]. The HMP was applied to solve the HOCPs with predefined shift sequences. The solution of the TPBVP and the initialization of the problem were not disclosed.

Remark: Determining the controls through minimizing instantaneous energy consumption is insufficient for this thesis, since it does not incorporate shift processes. The HMP from [26] provides inspiration but cannot be taken over, due to its requirement of predefined shift sequences. Furthermore, the algorithms for the solution process were not disclosed in the literature. It is important to analyse and reduce the time complexity of the algorithms for an online EMS.

3.3 Multi-Criteria Operation Strategies

A MCOS considers energy or fuel consumption and other aspects. The often-considered secondary aspects are discussed in the following paragraph with non-exhaustive examples. The realization of a MCOS is later discussed.

Many MCOSs for HEVs and PHEVs considered the pollutant emission, e.g. [15] and [124]. Some considered the noise of engines, e.g. [125] and [126]. Some considered the aging of battery, e.g. [127] and [128]. Some works tried to reduce the shift count by adding a penalty, so that the drivability was not compromised too much, e.g. [25] and [33].

The main approach to realize a MCOS is to formulate the OCP with a multi-criteria cost function

l that is defined as

$$l = l_{\text{energy}} + \beta \cdot l_{\text{2nd aspect}}, \quad (3.1)$$

where β is a weighting factor. The arguments of the cost functions are omitted for simplicity. Some secondary aspects are considered in the EMS by introducing extra constraints on states and controls. For instance the noise constraints was introduced in [125] and [129, pp. 115-122], because of the masking effect (see the references for more information). Another way to include a secondary aspect is to consider it as an endpoint functional in the cost functional as

$$J = m_{\text{2nd aspect}} + \int_{t_0}^{t_f} l_{\text{energy}} dt, \quad (3.2)$$

where J is the multi-criteria functional and $m_{\text{2nd aspect}}$ is the endpoint functional that evaluates the continuous states and the discrete state at the final time point. The arguments of the cost function and the functionals are omitted for simplicity.

3.4 Summary and Other Aspects

Summary

The online EMS in this thesis is constructed in an MPC framework. The research of the EMSs for HEVs and PHEVs have provided the evidence for the theoretical comparison of the solution methods in Section 2.4.4 from an application point of view. The HMP is to be embedded in the predictive EMS for its advantages in time complexity and close-to-optimum solution. This thesis needs to overcome its disadvantages regarding convergence and initialization by adopting a proper numerical method and developing an initialization strategy. Importantly, the method developed in this thesis shall cast no constraint on the shift sequences. After the review of EMSs for HEVs, PHEVs and BEVs, the EMS to be developed in this thesis is unseen in the literature. Additionally, this thesis is to provide the algorithms of the HMP solution method, which has not been properly disclosed in the literature.

For the reasons provided in Section 2.4.2, the minimization of the instantaneous Hamiltonian at (2.54) is important for the HMP in terms of accuracy and time complexity, but was not discussed in the literature. Other than the algorithm of the online EMS, this thesis will provide mathematical analysis of the instantaneous Hamiltonian with the goal of reducing time complexity.

A two-drive multi-speed BEV can be seen as an example of multi-drive multi-speed BEVs, in which more EMs and STs are coupled in the two-drive BEVs in Figure 3.1. It is worth studying,

how the developed EMS can be transferred to multi-drive multi-speed BEVs. Such an attempt is unseen.

A MCOS considering the service life of the powertrain is to be developed. The multi-criteria HOCP is to be formulated in such a way that the algorithms in the predictive EMS require little to none change. The service life is to be modelled and evaluated, so that the multi-criteria cost functional can be formulated accordingly.

Other aspects

The importance of a predictor for a predictive EMS has been stressed in Section 3.2.2.1. Different applications of speed prediction with various input information have been reviewed in [130]. Speed prediction is principally time series prediction, which may have been a dream of human since the dawn of time. Nowadays, the time series prediction is attempted mainly through machine learning and statistics. Books on this topic include [131] and [132]. A large range of methods were compared against each other using the M3-competition data in [133, 134] w.r.t. their accuracy, necessary preprocessing methods, computational requirements etc. This thesis chooses to use the Markov Chain model, a widely used statistical modelling method with proper accuracy and fairly few parameters⁴⁶, but not to invest a large amount of time into adopting and developing a neural network (NN) with a complicated architecture. The reasons are: 1) Advanced NNs under the category “deep learning”, especially the transformer model [136], have emerged in recent years and have led to the breakthrough in application in 2023, which makes it difficult to identify a state-of-the-art NN for time-series prediction; 2) On the other hand, the motivations and goals of this thesis are mainly relevant to the components “model” and “solver” in the predictive EMS shown in Figure 3.4, which makes a state-of-the-art prediction model not a prerequisite. However, the influence of the prediction model is not neglected. Section 6.3.4 investigates the influences of the accuracy of the prediction.

One may ask, if machine learning methods can replace the EMSs that are based on mathematical solution methods. According to Andrew Ng, for generating optimal control function, current machine learning methods still fall short of the performance of the mathematical solution methods [137]. Even if the situation would change one day, training machine learning methods might still require reference values generated by the EMS developed by mathematical solution methods.

⁴⁶ The statement is based on the comparative study [135], in which several methods chosen from [133, 134] are adopted to perform the acceleration and speed prediction with the driving data collected from a test vehicle of the Institute of Mechatronic Systems at the Technical University Darmstadt.

4 Hybrid System and Problem Formulation

In this chapter, a two-drive multi-speed powertrain called “Speed4E” is modelled as a hybrid system with proper consideration of its shift processes. Thereafter, a HOCP with a cost functional of energy consumption is formulated.

4.1 Speed4E Powertrain

Institute of Mechatronic System (IMS) is part of the research project Speed4E that desires to develop a high-speed two-drive multi-speed powertrain (see Section 4.1 for its topology). Beside combining the advantages of two-drive and multi speed powertrains, high-speed EMs can improve the energy density of the powertrain. Figure 4.1 qualitatively shows the benefit of the application of high-speed EM w.r.t. weight. Speed4E powertrain, consisting of EMs with a maximum speed of 50.000 rpm, is anticipated to reduce half of the mass compared to a Reference Design from [138] and to reduce 10 % of the mass of the Speed2E powertrain⁴⁷ developed in the proceeding project. The downsize makes Speed4E powertrain possible to be fit into the engine compartment of BMW i3s, a front-wheel drive test vehicle provided by Speed4E project partner BMW group.

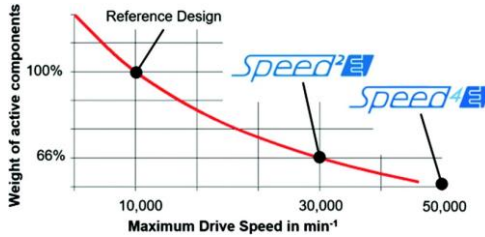


Figure 4.1: Advantages of high-speed e-drives w.r.t. weight of active components [139].

As shown in Figure 4.2, Speed4E powertrain consists of two EMs propelling the vehicle through two respective sub-transmissions (ST). The ST1 is a planetary gear transmission with a fixed ring gear. The ST2 is a three-stage two-speed spur gear transmission that enables three gear positions (G): the 1st, the 2nd and the neutral gear position. Their overall gear ratios can be calculated by sequential multiplications as

$$i_1 = i_{1,1} \cdot i_{1,2} \cdot i_{FD}, \quad (4.1)$$

⁴⁷ The project Speed2E developed a powertrain consisting of EMs with maximum speed of 30.000 rpm that was successfully validated on the testbench.

$$i_2 = \begin{cases} i_{2,1} \cdot i_{2,2_1} \cdot i_{2,3} \cdot i_{FD}, & G = 1, \\ i_{2,1} \cdot i_{2,2_2} \cdot i_{2,3} \cdot i_{FD}, & G = 2, \\ 0, & G = 3, \end{cases} \quad (4.2)$$

$$G \in \{1, 2, 3\}, \quad (4.3)$$

where $i_{1,1}$ denotes the gear ratio of the planetary gear and $i_{1,2}$ the gear ratio between the final drive and the carrier; i_{FD} denotes the gear ratio of the final drive; $i_{2,1}$, $i_{2,2}$ and $i_{2,3}$ denote the gear ratios of the first, the second and the third stage of the spur gear pairs respectively; $i_{2,2_1}$ and $i_{2,2_2}$ denote the gear ratios of the second stage in 1st and 2nd gear positions, respectively; when both gear positions are disengaged, namely $G = 3$, the ST2 is in the neutral position. Please note that its gear ratio is set to zero for a computational purpose, i.e. the output torque of the ST2 with $G = 3$ is simply zero. The reduced rotational inertia of the vehicle about the rotational axis of the wheels (I_{red}) is dependent on G . See Appendix A for its calculation and necessary parameters.

An innovative dual coil permanent excited linear actuator (LA) is developed to perform fast shifting between these gear positions [140]. The shift sleeve (SS) is moved by the LA to engage or disengage the dog clutches (DC) according to the requested gear position.

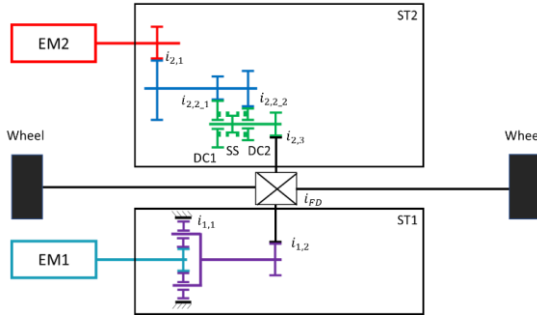


Figure 4.2: Topology of Speed4E powertrain [141]

4.2 Longitudinal Dynamics

The output torque of the powertrain overcomes the driving resistances and provides acceleration,

$$T_{total} = p(G, v, \dot{v}) = (mass_{eq}(G) \cdot \dot{v} + c_{roll}(v) \cdot mass \cdot g + 0.5 \cdot c_a \cdot \rho_{air} \cdot A_f \cdot v^2)r, \quad (4.4)$$

$$mass_{eq}(G) = mass + \frac{I_{red}(G)}{r^2}, \quad (4.5)$$

where T_{total} is the total output torque on the wheels that is determined by a function p with parameters the gear position G , the vehicle speed v and the acceleration \dot{v} . The vehicle is considered driving on an even ground so that slope resistance is omitted. In the context of EMSs, the function p can be used to determine the requested total torque $T_{\text{total,req}}$ based on the requested acceleration \dot{v}_{req} . $mass_{\text{eq}}$ is the equivalent mass, which considers $mass$, the vehicle mass, and I_{red} . The rolling resistance is calculated according to the coefficient c_{roll} and the gravity force normal to the ground. The aerodynamic resistance is calculated according to the coefficient c_a , the air density ρ_{air} , the frontal area A_f and v . r is the dynamic radius of the wheel. See Appendix A for their values.

If the total output torque is given, the acceleration can be determined by

$$\dot{v} = b(G, v, T_{\text{total}}) = \frac{T_{\text{total}} - c_{\text{roll}}(v) \cdot mass \cdot g - 0.5 \cdot c_a \cdot \rho_{\text{air}} \cdot A_f \cdot v^2}{mass_{\text{eq}}(G)}. \quad (4.6)$$

The total output torque is provided by both EMs

$$T_{\text{total}} = T_{\text{ST1}} \cdot i_1 + T_{\text{ST2}} \cdot i_2(G), \quad (4.7)$$

$$T_{\text{ST1}} = \phi_1(\omega_1, T_1) = T_1 - \frac{PL_{\text{ST1}}(\omega_1, T_1)}{\omega_1}, \quad (4.8)$$

$$T_{\text{ST2}} = \phi_{2,G}(\omega_2, T_2) = \begin{cases} T_2 - \frac{PL_{\text{ST2,G}}(\omega_2, T_2)}{\omega_2}, & \text{if } G = 1 \text{ or } 2, \\ 0, & \text{if } G = 3, \end{cases} \quad (4.9)$$

where T_{ST} denotes the effective torque of a ST, i.e. the input torque having the torque loss deducted, which is described by the function $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$. ω and T denote the angular velocity and the torque of an EM. i is the gear ratio, which is gear position dependent in the case of the ST2. Power losses PL are modelled as look-up tables (see Appendix B) based on the simulation results from the project partners (see Appendix B). By performing interpolation, the function $\phi': \mathbb{R}^2 \rightarrow \mathbb{R}$ that determines the input torque based on the effective torque can be numerically acquired, as

$$T_1 = \phi'_1(\omega_1, T_{\text{ST1}}), \quad (4.10)$$

$$T_2 = \phi'_{2,G}(\omega_2, T_{\text{ST2}}). \quad (4.11)$$

The function ϕ is formally examined in Appendix D.

The total electric power includes the mechanical powers and the power losses $PL_{\text{EM\&PE}}$ that sum up those of EMs and PEs that are modelled as look-up tables (see Appendix B),

$$P_{ele} = \omega_1 \cdot T_1 + PL_{EM\&PE1}(\omega_1, T_1) + \omega_2 \cdot T_2 + PL_{EM\&PE2}(\omega_2, T_2). \quad (4.12)$$

4.3 Shift Processes

The angular position control (APC) is the core of the shift processes. The main idea is to use the EM2 to regulate the relative angular position between the DC and the SS, so that a) the friction between the teeth of the SS and those of the DC during the DC disengagement does not occur, and b) teeth-to-teeth situations during the DC engagement do not occur [141]. This subsection describes the shift processes modelled for the use of the EMS based on the experiment and simulation results that was presented in [142].

Figure 4.3 shows a shift process used in the EMS from the 1st to the 2nd gear position with an initial speed of 50 km/h and an acceleration of 3.6 m/s². Before the APC takes place, T_{EM1} ramps up to take over all propelling torque, while T_{EM2} ramps down towards zero in the “Balance” phase, in which the total output torque on the wheels remains constant. In this specific example, T_{EM1} reaches its maximum before T_{EM2} reaches zero, after which T_{EM1} remains at its maximum throughout the whole shift process and T_{EM2} further ramps down to zero in “Decrease” phase, as the total output torque on the wheels and the vehicle acceleration (blue curve) decrease. As soon as T_{EM2} is

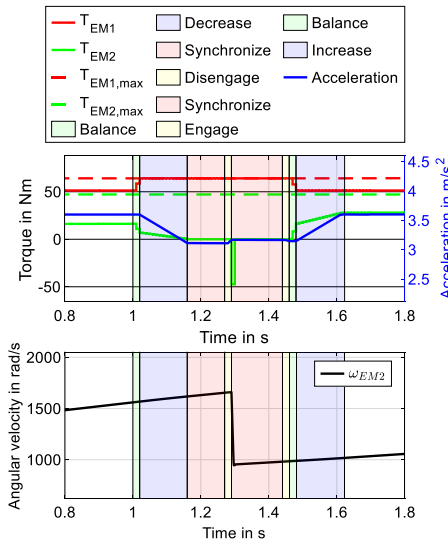


Figure 4.3: Shift process from 1st to 2nd gear position in the EMS

zero, APC starts with the first “Synchronize” phase, during which T_{EM2} is regulated between 0 Nm and 0.05 Nm in the real-world application [142] to minimize the friction between the SS and the DC. In the model for EMS, T_{EM2} is simplified to be zero in this phase. Thereafter, the SS is disengaged from the DC1, which shifts the ST2 into its neutral position. In the second “Synchronize” phase, ω_2 is firstly decelerated towards the new value in the 2nd gear position, i.e. $\omega_{2,target} = v/r \cdot i_2(G_{target})$. The angular position of the EM2 is then regulated, through T_{EM2} with small value [142], to ensure no teeth-to-teeth situation in the “Engage” phase. In the model for EMS, T_{EM2} is firstly its minimum during deceleration then simplified to be zero during APC (green solid curve). After the APC finishes, i.e. the ST2 being shifted into the 2nd gear position, T_{EM1} and T_{EM2} ramp to their new control values determined by the EMS with the second “Balance” and “Increase” phases in a mirror manner to the first “Balance” and “Decrease” phases. During the APC, minor changes of the acceleration can be noticed, since the change of gear position causes the change of rotational inertia of the powertrain and therefore the equivalent mass m_{eq} .

Figure 4.4(a) and (b) show a shift process from 1st to neutral gear position and a shift process from neutral to 1st gear position with respective speed and acceleration conditions. Shifting into or out of the neutral position can be considered the first half, namely from the first “Balance” phase to the “Disengage” phase, or the second half, namely the second “Synchronize” phase to the end,

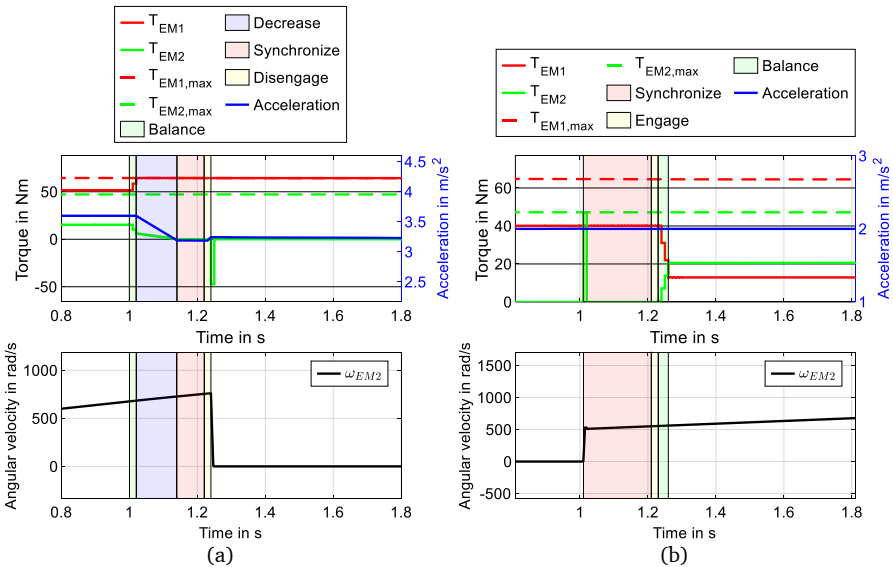


Figure 4.4: Shift process in the EMS. (a): from 1st to neutral gear position. (b): from neutral to 1st gear position.

of the shift process between 1st and 2nd gear position. The differences is that the EM2 decelerates to standstill or accelerates from standstill without further synchronization or previous synchronization, respectively.

Table 4.1. Durations of different phases in a shift process

| Phase | Duration |
|----------------------|--|
| First “Synchronize” | 0.1 s |
| Second “Synchronize” | velocity and shift process dependent look-up table |
| “Disengage” | 0.02 s |
| “Engage” | 0.02 s |

The simplifications of the shift processes used in the EMS are: 1) The change of torque is considered linear. 2) During synchronization, the torque of EM2 is considered zero, though it’s a minor value close to zero to regulate the angular position. 3) When the angular velocity of EM2 is accelerated or decelerated, the torque jumps to its maximum or minimum without ramps. 4) The durations of the “Disengage”, “Engage” and the “Synchronize” phases are simplified to be constant based on the experiments from [141] (see Table 4.1), while the duration of the second “Synchronize” phase, including deceleration and APC, is modelled as a vehicle speed and shift process dependent look-up table based on the simulation results [142].

4.4 Hybrid System Formulation

Speed4E Vehicle described so far is modelled as a hybrid system $\mathbb{H}_{\text{Speed4E}}$ according to **Definition 2.4**. The continuous states consist of the angular velocities of EM1 and EM2 $\mathbf{x} := [\omega_1, \omega_2]^T$. The torques of both EM are the continuous controls $\mathbf{u} := [u_1, u_2]^T = [T_1, T_2]^T$. Figure 4.5 shows the hybrid automaton of $\mathbb{H}_{\text{Speed4E}}$ with the discrete states q_1, q_2 and q_3 represent the 1st, 2nd and neutral gear positions, and q_4, q_5, \dots, q_9 represent the shift processes among different gear positions.

The admissible discrete controls for q_1, q_2 and q_3 are summarized in Table 4.2. They cause controlled switchings, which are marked with blue color in Figure 4.5. It is considered equivalent to write $q = q_1, q_2, \dots, q_9$ and $q = 1, 2, \dots, 9$, respectively. Each row represents a discrete set \mathcal{B}_q , from which an admissible discrete control ϖ_q is chosen. For instance, when the current discrete state is q_1 , the discrete control can command the system to remain q_1 or to enter either q_4 (shift from the 1st to the 2nd gear position) or q_6 (shift from the 1st to the neutral gear position), i.e. $\mathcal{B}_{q_1} := \{q_1, q_4, q_6\}$. A shift process is divided into several phases. They are called sub-states for differentiation from the discrete states $\{q_1, q_2, \dots, q_9\}$. When a discrete state representing a shift process is selected, the discrete control switches among the corresponding sub-states sequentially according

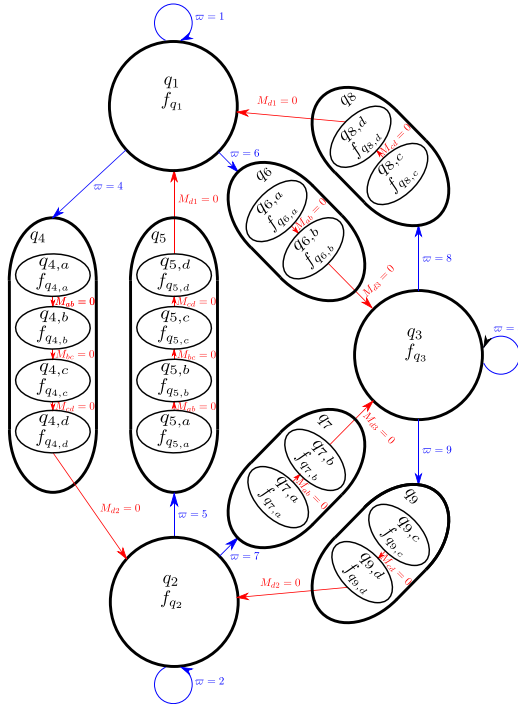


Figure 4.5: Hybrid automaton of Speed4E powertrain

to the corresponding switching manifolds M (marked red). Table 4.3 lists the autonomous admissible discrete controls. Note that the possible discrete control is singular at any discrete state or sub-state. It is considered equivalent to write $\varpi = q_1, q_2, q_3$ and $\varpi = 1, 2, 3$, respectively. The discrete set of $\mathbb{H}_{\text{Speed4E}}$ includes discrete states and sub-states $\mathcal{Q} := \{q_1, q_2, \dots, q_9, q_{4,a}, \dots, q_{4,d}, \dots, q_{9,d}\}$. The discrete transition function is simply defined as

$$q(t^+) = \varpi(t). \quad (4.13)$$

Table 4.2. Admissible discrete control set \mathcal{B} , controlled switching

| q | Admissible discrete control (\mathcal{B}_q) | | |
|-----|---|---|---|
| 1 | 1 | 4 | 6 |
| 2 | 2 | 5 | 7 |
| 3 | 3 | 8 | 9 |

Table 4.3. Admissible discrete control set \mathcal{B} , autonomous switching

| q | \mathcal{B}_q | q | \mathcal{B}_q | q | \mathcal{B}_q | q | \mathcal{B}_q | q | \mathcal{B}_q | q | \mathcal{B}_q |
|-----|-----------------|-----|-----------------|-----|-----------------|-----|-----------------|-----|-----------------|-----|-----------------|
| 4 | 4,a | 5 | 5,a | 6 | 6,a | 7 | 7,a | 8 | 8,c | 9 | 9,c |
| 4,a | 4,b | 5,a | 5,b | 6,a | 6,b | 7,a | 7,b | 8,c | 8,d | 9,c | 9,d |
| 4,b | 4,c | 5,b | 5,c | 6,b | 6,c | 7,b | 7,c | 8,d | 1 | 9,d | 2 |
| 4,c | 4,d | 5,c | 5,d | 6,c | 6,d | 7,c | 7,d | - | - | - | - |
| 4,d | 2 | 5,d | 1 | 6,d | 3 | 7,d | 3 | - | - | - | - |

The following elaborates the dynamics of the hybrid system.

4.4.1 Discrete States q_1, q_2 and q_3

The ST2 stays in the 1st, 2nd or neutral position, respectively. The indexed system dynamics can be derived from (4.6) as

$$\begin{aligned} \dot{x}_q &= f_q(x, T_{\text{total}}) \\ &= \frac{b(G, v, T_{\text{total}})}{r} \cdot \begin{pmatrix} i_1 \\ i_2(G) \end{pmatrix}, \quad \forall t \in [t_{0,q}, t_{f,q}], \end{aligned} \quad (4.14)$$

$$T_{\text{total}} = \phi_1(\omega_1, T_1) \cdot i_1 + \phi_{2,G}(\omega_2, T_2) \cdot i_2(G), \quad (4.15)$$

$$v = \frac{\omega_1}{i_1} \cdot r, \quad (4.16)$$

$$G = q, \quad (4.17)$$

$$x(t_{0,q}) = x_{0,q}, \quad (4.18)$$

where $t_{0,q}$ and $t_{f,q}$ denote the start and end time of q ; the initial state $x_{0,q}$ is given. Since $q = q_1, q_2, q_3$ and $q = 1, 2, 3$ are defined equivalent, there is, for instance, $G = q_1 = 1$.

4.4.2 Discrete States q_4, q_5, \dots, q_9

Table 4.4 shows the representation of each sub-state in a SP. q_4 and q_5 represent the shift processes from the 1st to 2nd gear position and its reversed order, in which all four sub-states take place. q_6 and q_7 represent the shift processes from the 1st and 2nd gear position to the neutral position, respectively. As Figure 4.4 shows, they consist of sub-states a and b . q_8 and q_9 represent the shift processes from the neutral position to the 1st and 2nd gear position, respectively. As Figure 4.4(b) shows, they consist of sub-states c and d . Regardless of sub-states, the continuous control during a shift process is predefined. Therefore, the dynamics of $\dot{x}_q, q \in \mathcal{Q} \setminus \{q_1, q_2, q_3\}$ are autonomous. “\” indicates “exclude”.

Table 4.4. States in a shift process

| Sub-states | Phases in a shift process |
|------------|-------------------------------------|
| $q_{j,a}$ | “Balance” and “Decrease” |
| $q_{j,b}$ | First “Synchronize” and “Disengage” |
| $q_{j,c}$ | Second “Synchronize” and “Engage” |
| $q_{j,d}$ | “Increase” and second “Balance” |

Shift processes can be divided into four categories that are the combinations of two conditions: 1) positive or negative output torque on the wheels, i.e. $T_{\text{total}} > 0$ or i.e. $T_{\text{total}} \leq 0$ and 2) upshift or downshift, i.e. $\omega_{2,\text{target}} < \omega_2$ or $\omega_{2,\text{target}} > \omega_2$. The following subsections (from 4.4.2.2 to 4.4.2.4) describe the mathematical model of each sub-state in the case of a upshift with positive output torque, i.e. $T_{\text{total}} > 0$ and $\omega_{2,\text{target}} < \omega_2$. The changes required for the shift processes in the rest three categories are shown in 4.4.3.

4.4.2.1 “Balance” and “Decrease” phases

As discussed in Section 4.3, the duration of the “Balance” phase is the smaller one between the duration for T_1 to reach its maximum ($\tau_{1,\text{Fade}}$) and the duration for T_2 to reach zero ($\tau_{2,\text{Fade}}$). They are determined by

$$\tau_{1,\text{Fade}} = \frac{(T_{1,\text{max}}(\omega_1) - T_1(t_{0,j})) \cdot i_1}{TS_{\text{Bal}}}, \quad (4.19)$$

$$\tau_{2,\text{Fade}} = \frac{(T_2(t_{0,q})) \cdot i_2(q(t_{0,j}^-))}{TS_{\text{Bal}}}, \quad (4.20)$$

$$\tau_{\text{Bal}} = \min(\tau_{1,\text{Fade}}, \tau_{2,\text{Fade}}), \quad (4.21)$$

where TS_{Bal} denotes the rate of the change of output torque on the wheels in the “Balance” phase; $t_{0,j}$ the time point when discrete state enters q_j .

A “Decrease” phase takes place, if T_1 reaches its maximum before T_2 reaches zero. The duration of the “Decrease” phase is

$$\tau_{\text{Decr}} = \begin{cases} \frac{T_2(t_{0,j} + \tau_{\text{Bal}}) \cdot i_2(q(t_{0,j}^-))}{TS_{\text{Decr}}}, & \tau_{\text{Bal}} < \tau_{2,\text{Fade}}, \\ 0, & \tau_{\text{Bal}} = \tau_{2,\text{Fade}}, \end{cases} \quad (4.22)$$

$$\tau_a = \tau_{\text{Bal}} + \tau_{\text{Decr}}, \quad (4.23)$$

where TS_{Decr} denotes the slope of the change of output torque on the wheels in the “Decrease” phase. If the duration of the “Balance” phase equals $\tau_{2,\text{Fade}}$, the “Decrease” phase does not take place, which means τ_{Decr} equals zero. The duration of the sub-state $q_{j,a}$ (τ_a) is the sum of the duration of the “Balance” and “Decrease” phases. In this sub-state, the torque variables in \mathbf{u} are determined by

$$T_1 = \begin{cases} T_1(t_{0,q}) + \frac{TS_{\text{Bal}}}{i_1} \cdot (t - t_{0,j}), & t \in (t_{0,j}, t_{0,j} + \tau_{\text{Bal}}], \\ T_{1,\text{max}}(\omega_1), & t \in (t_{0,j} + \tau_{\text{Bal}}, t_{0,j} + \tau_a], \end{cases} \quad (4.24)$$

$$T_2 = \begin{cases} \phi'_{2,G} \left(\omega_2, \frac{T_{\text{total}} - \phi_1(\omega_1, T_1) \cdot i_1}{i_2 (q(t_{0,j}^-))} \right), & t \in (t_{0,j}, t_{0,j} + \tau_{\text{Bal}}], \\ T_2(t_{0,j} + \tau_{\text{Bal}}) - \frac{TS_{\text{Decr}}}{i_2 (q(t_{0,j}^-))} \cdot (t - t_{0,j} - \tau_{\text{Bal}}), & t \in (t_{0,j} + \tau_{\text{Bal}}, t_{0,j} + \tau_a], \end{cases} \quad (4.25)$$

where T_1 fades to its maximum with a slope dependent on TS_{Bal} , while T_2 fades to zero firstly compensating the change of T_1 and secondly dependent on TS_{Decr} , if necessary. Since the DC is not disengaged, (4.15) holds with

$$G = q(t_{0,j}^-), \quad t \in (t_{0,j}, t_{0,j} + \tau_a]. \quad (4.26)$$

An autonomous transition from state $q_{j,a}$ to $q_{j,b}$ happens, when the time-dependent switching manifold $M_{ab}(t) = 0$, which is defined as

$$M_{ab}(t) = t - t_{0,j} - \tau_a. \quad (4.27)$$

4.4.2.2 First “Synchronize” and “Disengage” phases

In $q_{j,b}$, the first “Synchronize” phase takes place, followed by the “Disengage” phase. Both phases are assumed to last for constant time, as discussed in Section 4.3. The duration of this sub-state (τ_b) and the torques are

$$\tau_b = \tau_{\text{syn1}} + \tau_{\text{Disengage}} \quad (4.28)$$

$$T_1 = \frac{T_{\text{total}}(q(t_{0,j}^-)) - \phi_{2,q}(t_{0,j}^-)(\omega_2, 0) \cdot i_2(q(t_{0,j}^-))}{i_1}, \quad t \in (t_{0,j,b}, t_{0,j,b} + \tau_b], \quad (4.29)$$

$$T_2 = 0, \quad t \in (t_{0,j,b}, t_{0,j,b} + \tau_b], \quad (4.30)$$

where EM1 propels the vehicle alone, also compensating the drag torque of ST2. $t_{0,j,b}$ is the time point when the sub-state $q_{j,b}$ is activated. Since the DC is not disengaged, (4.14)-(4.18) hold with

$$G = q(t_{0,j}^-), \quad t \in (t_{0,j,b}, t_{0,j,b} + \tau_b]. \quad (4.31)$$

An autonomous transition from sub-state $q_{j,b}$ to $q_{j,c}$ happens, when the time-dependent switching manifold $M_{bc}(t) = 0$, which is defined as

$$M_{bc}(t) = t - t_{0,j,b} - \tau_b. \quad (4.32)$$

4.4.2.3 Second ‘‘Synchronize’’ and ‘‘Engage’’ phases

In $q_{j,c}$, the second ‘‘Synchronize’’ phase with a duration τ_{syn2} takes place, followed by the ‘‘Engage’’ phase with constant duration, as discussed in Section 4.3. τ_{syn2} depends on the vehicle speed and the shift process q_j . It includes a) the duration for EM2 to accelerate τ_{acc} , negative acceleration in this case, and b) the duration of the APC τ_{APC} . The duration of this sub-state (τ_c) and the torques are

$$\tau_c = \tau_{\text{acc}} + \tau_{\text{APC}} + \tau_{\text{Engage}} = \tau_{\text{syn2},q_j}(v) + \tau_{\text{Engage}}, \quad (4.33)$$

$$T_1 = \min\left(T_{1,\text{max}}(\omega_1), \phi'_1\left(\omega_1, \frac{T_{\text{total}}(G)}{i_1}\right)\right), \quad t \in (t_{0,j,c}, t_{0,j,c} + \tau_c], \quad (4.34)$$

$$T_2 = 0, \quad t \in (t_{0,j,c}, t_{0,j,c} + \tau_c], \quad (4.35)$$

$$G = 3, \quad t \in (t_{0,j,c}, t_{0,j,c} + \tau_c], \quad (4.36)$$

where the gear position is neutral, since the DC is disengaged. The dynamics follow

$$\begin{aligned} \dot{x}_q &= f_q(x) \\ &= \left(\begin{array}{c} \frac{b(G, v, \phi_1(\omega_1, T_1) \cdot i_1)}{r} \\ \frac{T_2}{I_{\text{ST2,input}}} \end{array} \cdot i_1 \right), \quad \forall t \in [t_{0,q}, t_{f,q}], \end{aligned} \quad (4.37)$$

$$v = \frac{\omega_1}{i_1} \cdot r, \quad (4.38)$$

$$T_2 = \begin{cases} T_{2,\min}, & t \in (t_{0,j,c}, t_{0,j,c} + \tau_{\text{acc}}], \\ 0, & t \in (t_{0,j,c} + \tau_{\text{acc}}, t_{0,j,c} + \tau_c], \end{cases} \quad (4.39)$$

$$\tau_{\text{acc}} = \frac{\left(\frac{\omega_1}{i_1} \cdot i_2 (\varpi_{q_{j,d}}) - \omega_2(t_{0,j,c})\right) I_{\text{ST2,input}}}{T_2}, \quad (4.40)$$

where $I_{\text{ST2,input}}$ is the rotational inertia around the EM2 rotor axis including the rotor, the first and the second stages of the ST2. For $q_j \in \{q_4, q_5, \dots, q_9\}$, $\varpi_{q_{j,d}}$ is equivalent to the target gear position (see Table 4.3).

An autonomous transition from the sub-state $q_{j,c}$ to $q_{j,d}$ happens, when the time-dependent switching manifold $M_{cd}(t) = 0$, which is defined as

$$M_{cd}(t) = t - t_{0,j,c} - \tau_c. \quad (4.41)$$

4.4.2.4 "Increase" and second "Balance" phases

In $q_{j,d}$, both torques fade to $T_{1,\text{new}}$ and $T_{2,\text{new}}$, respectively, which are determined by the EMS for the new gear position. Similar to the sub-state $q_{j,a}$, both "Balance" and "Increase" phases might take place. Their durations are determined by

$$\tau_{1,\text{Fade}} = \frac{(T_1(t_{0,j,d}) - T_{1,\text{new}}) \cdot i_1}{TS_{\text{Bal}}}, \quad (4.42)$$

$$\tau_{2,\text{Fade}} = \frac{(T_{2,\text{new}} - T_2(t_{0,j,d})) \cdot i_2 (\varpi_{q_{j,d}})}{TS_{\text{Bal}}}, \quad (4.43)$$

$$\tau_{\text{Bal}} = \min(\tau_{1,\text{Fade}}, \tau_{2,\text{Fade}}), \quad (4.44)$$

$$\tau_{\text{Incr}} = \begin{cases} \frac{(T_{2,\text{new}} - T_2(t_{0,j,d} + \tau_{\text{Bal}})) \cdot i_2 (\varpi_{q_{j,d}})}{TS_{\text{Incr}}}, & \tau_{\text{Bal}} < \tau_{2,\text{Fade}}, \\ 0, & \tau_{\text{Bal}} = \tau_{2,\text{Fade}}, \end{cases} \quad (4.45)$$

$$\tau_d = \tau_{\text{Bal}} + \tau_{\text{Incr}}, \quad (4.46)$$

where $\varpi_{q_{j,d}}$ is equivalent to the target gear (see Table 4.3). In this sub-state, T_1 is determined by

$$T_1 = \begin{cases} T_1(t_{0,j,d}) + \frac{TS_{\text{Bal}}}{i_1} \cdot (t - t_{0,j,d}), & t \in (t_{0,j,d}, t_{0,j,d} + \tau_{\text{Bal}}], \\ T_{1,\text{new}}, & t \in (t_{0,j,d} + \tau_{\text{Bal}}, t_{0,j,d} + \tau_d], \end{cases} \quad (4.47)$$

$$T_2 = \begin{cases} \phi'_{2,G} \left(\omega_2, \frac{T_{\text{total}} - \phi_1(\omega_1, T_1) \cdot i_1}{i_2(\varpi_{q_{j,d}})} \right), & t \in (t_{0,j,d}, t_{0,j,d} + \tau_{\text{Bal}}], \\ T_2(t_{0,j,d} + \tau_{\text{Bal}}) + \frac{TS_{\text{Incr}}}{i_2(\varpi_{q_{j,d}})} \cdot (t - t_{0,j,d} - \tau_{\text{Bal}}), & t \in (t_{0,j,d} + \tau_{\text{Bal}}, t_{0,j,d} + \tau_d]. \end{cases} \quad (4.48)$$

Since the DC of the new gear position is engaged, (4.14)-(4.18) hold with

$$G = \varpi_{q_{j,d}}, \quad t \in (t_{0,j,d}, t_{0,j,d} + \tau_d]. \quad (4.49)$$

Since $\varpi = q_1, q_2, q_3$ and $\varpi = 1, 2, 3$ are defined equivalent, there is, for instance, $G = q_1 = 1$. An autonomous transition from the sub-state $q_{j,d}$ to the next discrete state $\varpi_{q_{j,d}}$ happens, when the time-dependent switching manifold $M_{j\varpi_{q_j}}(t) = 0$, which is defined as

$$M_{j\varpi_{q_j}}(t) = t - t_{0,j,d} - \tau_d. \quad (4.50)$$

4.4.3 Shift Process in Other Categories

In the case of $T_{\text{total}} \leq 0$, replace $T_{1,\text{max}}$ with $T_{1,\text{min}}$ in (4.19), (4.24) and (4.34). In the case of $\omega_{2,\text{target}} > \omega_2$, replace $T_{2,\text{min}}$ with $T_{2,\text{max}}$ in (4.39).

4.5 Hybrid Optimal Control Problem of Minimal Energy Consumption

On a given time interval $[t_0, t_f]$, Speed4E vehicle is controlled by an EMS to minimize its energy consumption. The torques are the continuous controls, i.e. $\mathbf{u} = (u_1, u_2)^T = (T_1, T_2)^T$, of the system \mathbb{H} modelled in Section 4.4. The problem is formulated as a HOCP defined by **Definition 2.12** as

$$(\varpi^*(\cdot), \mathbf{u}^*(\cdot)) = \underset{\varpi(\cdot) \in \mathcal{B}_{q(\cdot)}, \mathbf{u}(\cdot) \in \mathcal{U}(\cdot)}{\text{argmin}} J(q(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)), \quad (4.51)$$

$$J(q(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)) = m(\mathbf{x}(t_{f,\text{B}})) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (4.52)$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + PL_{EM1}(\omega_1, u_1) + PL_{EM2}(\omega_2, u_2), \quad (4.53)$$

$$m(\mathbf{x}(t_f)) = \beta \left\| \mathbf{x}(t_f) - \mathbf{x}_{t_f} \right\|^2, \quad (4.54)$$

$$\mathbf{x}_{t_f} = \frac{v_{\text{cycle}}(t_f)}{r} \cdot \left(i_1, i_2 \left(q(t_f) \right) \right)^T, \quad (4.55)$$

where the cost function (4.53) determines the electric power based on (4.12), which is not indexed (discrete state dependent), since the calculation of electric power based on the variables of the EMs does not require gear position. $\|\cdot\|$ calculates the Euclidean distance of a vector. The endpoint functional m imposes the endpoint constraint to the optimization, i.e. the angular velocities of both EM should correspond the speed from the driving cycle (v_{cycle}) and the gear position. β is a coefficient in the endpoint functional, whose value in the EMS is discussed in Section 5.2.1. The optimization subject to the hybrid system dynamics elaborated in Section 4.4 and the admissible discrete control set \mathcal{B}_q of the hybrid system \mathbb{H} . Additionally, following constraints need to be met,

$$q(t_0) = q_0, \quad (4.56)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad (4.57)$$

$$\mathbf{x}(t_f) = \mathbf{x}_{t_f}, \quad (4.58)$$

$$u_1 \in [T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)], \quad (4.59)$$

$$u_2 \in [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)], \quad (4.60)$$

$$\omega_1 \in [\omega_{1,\min}, \omega_{1,\max}], \quad (4.61)$$

$$\omega_2 \in [\omega_{2,\min}, \omega_{2,\max}], \quad (4.62)$$

$$q \in \mathcal{Q}. \quad (4.63)$$

where subscripts “0” and “ f ” indicate initial and final time point. Subscripts “min” and “max” indicate lower and upper bounds.

Please be aware that u and T are used interchangeably in the rest of the thesis.

4.6 Summary

In this chapter, hybrid system formulation in the context of HOCs and EMSs for a two-drive multi-speed powertrain is presented. The discussion focuses on a specific example, Speed4E powertrain, whose shift processes are represented in the hybrid system formulation by the corresponding discrete states. With the goal of reducing computational intensity, certain simplifications are introduced in such a way that the shift processes are primarily time dependent instead of solely continuous states dependent. The latter approach requires a small time step and a small convergence

tolerance, which are luxuries for algorithms developed for online operations. The duration parameters in the time dependent model are based on the simulation and experiment data from [141] and [142]. The HOCP for Speed4E powertrain is formulated based on the mathematical model of Speed4E powertrain.

5 Development of Energy Management Strategy

The hybrid model $\mathbb{H}_{\text{Speed4E}}$ has been developed in Chapter 4. This Chapter develops the predictive EMS embedded with the HMP solution method, whose structure is shown in Figure 5.1. There are two main components “Speed predictor” and “HMP solution” to be developed.

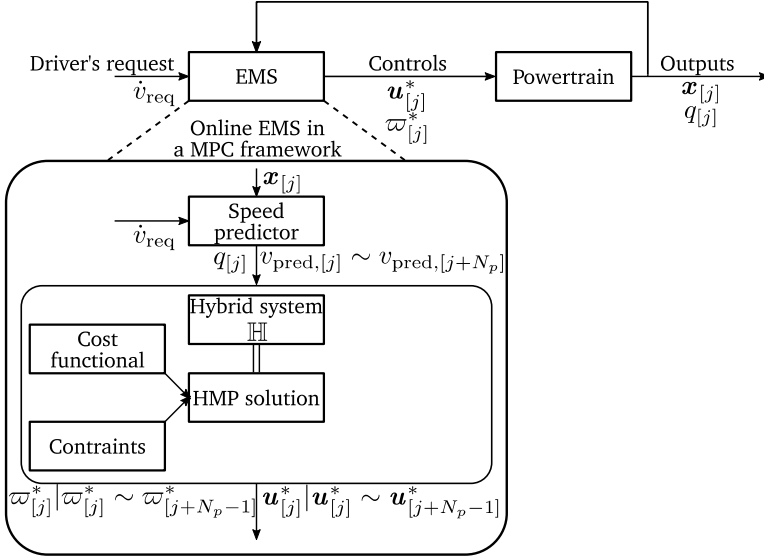


Figure 5.1: Basic structure of the predictive EMS embedded with the HMP solution method.

In the above illustrated control loop, the outputs of the powertrain are identical to the states (\mathbf{x}). At the time $t_{[j]}$, the driver requested acceleration (\dot{v}_{req}) and the current continuous states $\mathbf{x}_{[j]}$ are the inputs of the speed predictor to generate the predicted speed in the prediction horizon ($v_{\text{pred},[j]} \sim v_{\text{pred},[j+N_p]}$). Given the predicted speed, the current discrete state ($q_{[j]}$) and the hybrid system \mathbb{H} , the HMP solution method determines the optimal continuous control function ($\mathbf{u}_{\text{pred},[j]} \sim \mathbf{u}_{\text{pred},[j+N_p-1]}$) and the optimal discrete control sequence ($\omega_{\text{pred},[j]} \sim \omega_{\text{pred},[j+N_p-1]}$). See Section 3.2.2.1 for “ $N_p - 1$ ”. Their values in the first time step are provided to the powertrain.

5.1 Vehicle Speed Prediction

A Markov chain (MC) model predicts the speed in the prediction horizon v_{pred} . It's an important method to model stochastic processes on discrete time [102, pp. 587-601]. Let $\{mc(n), n = 0, 1, 2, \dots\}$ be an infinite sequence in a discrete finite state space. It is a MC if the probability of

moving to the next state depends only on the current state. A Transition Probability Matrix (TPM) can define the probability between MC states by

$$\text{tpm}_{ij} = P(mc(k+1) = mc_j | mc(k) = mc_i), \quad (5.1)$$

where tpm_{ij} denotes the i -th row and j -th column in the TPM and k the discrete time index. i and j index the MC states.

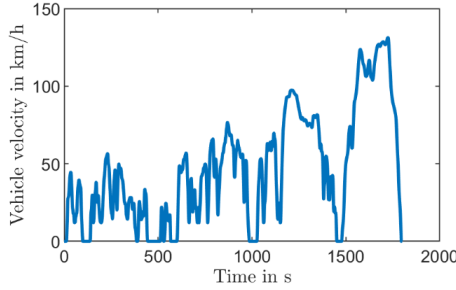


Figure 5.2: WLTC speed profile

Given a speed profile such as WLTC in Figure 5.2, a two-dimensional grid is generated by discretizing speed v (0 to 36.5 m/s, interval 0.75 m/s) and acceleration \dot{v} (-1.6 to 1.6 m/s², interval 0.032 m/s²). The speed and acceleration at each second are assigned to a point on the grid by nearest-neighbour method. A point on the grid is denoted by indexed MC state mc_i , which represents a pair (v, \dot{v}) . The TPM can be estimated as

$$\text{tpm}_{ij} = \frac{\text{count}(mc(k+1) = mc_j | mc(k) = mc_i)}{\text{count}(mc_i)}, \quad (5.2)$$

where “count(·)” counts the occurrences of the event in the argument. The MC states over the prediction horizon that is discretized into N_t time steps are sequentially

$$j = \max_j(\text{tpm}_{ij}), \quad mc(k) = mc_i, \quad (5.3)$$

$$mc(k+1) = mc_j, \quad k = 0, 1, \dots, N_t. \quad (5.4)$$

The speed at each time point can be calculated according to the grid assignment. In what follows, the procedure to generate predicted speed with the MC model is referred to as **predMC** with the current states, the requested acceleration and the length of prediction horizon as its inputs.

5.2 Hybrid Minimum Principle Solution Method

In this subsection, the general HOCBP in Section 4.5 in the prediction horizon is formulated, after which the specific formulation of the optimality conditions derived from HMP discussed in Section 2.4.1 is given. The TPBVP that is transformed from the HOCBP is provided, which is followed by its numerical solution method. All algorithms are presented.

5.2.1 Hybrid Optimal Control Problem in the Prediction Horizon

The torques are the continuous controls, i.e. $\mathbf{u} = (u_1, u_2)^T = (T_1, T_2)^T$, of the system $\mathbb{H}_{\text{Speed4E}}$. The HOCBP (4.51) is solved in the prediction horizon $[t_{0,p}, t_{f,p}]$, which is formulated as

$$(\varpi^*(\cdot), \mathbf{u}^*(\cdot)) = \underset{\varpi(\cdot) \in \mathcal{B}_{q(\cdot)}, \mathbf{u}(\cdot) \in \mathcal{U}(\cdot)}{\operatorname{argmin}} J(q(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)), \quad (5.5)$$

$$J(q(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)) = m(\mathbf{x}(t_{f,p})) + \int_{t_{0,p}}^{t_{f,p}} l(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (5.6)$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + PL_{EM1}(\omega_1, u_1) + PL_{EM2}(\omega_2, u_2), \quad (5.7)$$

$$m(\mathbf{x}(t_{f,p})) = \beta \left\| \mathbf{x}(t_{f,p}) - \mathbf{x}_{t_{f,p}} \right\|^2, \quad (5.8)$$

$$\mathbf{x}_{t_{f,p}} = \frac{v_{\text{pred}}(t_{f,p})}{r} \cdot \left(i_1, i_2 \left(q(t_{f,p}) \right) \right)^T, \quad (5.9)$$

subject to

$$q(t_{0,p}) = q_{0,p}, \quad (5.10)$$

$$\mathbf{x}(t_{0,p}) = \mathbf{x}_{0,p}, \quad (5.11)$$

$$\mathbf{x}(t_{f,p}) = \mathbf{x}_{t_{f,p}}, \quad (5.12)$$

$$u_1 \in [T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)], \quad (5.13)$$

$$u_2 \in [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)], \quad (5.14)$$

$$\omega_1 \in [\omega_{1,\min}, \omega_{1,\max}], \quad (5.15)$$

$$\omega_2 \in [\omega_{2,\min}, \omega_{2,\max}], \quad (5.16)$$

$$q \in \mathcal{Q}. \quad (5.17)$$

The algorithm **predictiveEMS** describes the predictive EMS. The function **solveHMP** solves the HOCBP with the HMP that is elaborated in the following sections. The square brackets at the subscripts convey the idea that the numerical calculations are performed in discrete time. [0] indicates the current time point. Based on $\mathbf{x}_{[0]}$ and the requested acceleration \dot{v}_{req} , **predMC** provides the unconstrained predicted speed profile $\mathbf{V}' := [v'_{\text{pred},[0]}, v'_{\text{pred},[1]}, \dots, v'_{\text{pred},[N_p]}]$ and the predicted acceleration $\dot{\mathbf{V}}' := [v'_{\text{pred},[0]}, \dot{v}'_{\text{pred},[1]}, \dots, \dot{v}'_{\text{pred},[N_p]}]$. β in (5.8) is parameterized to make sure that $\mathbf{x}(t_{f,p})$

from the solution correspond to the predicted speed, which is discussed in Section 6.3.1. It is necessary to have the solution process less sensitive to (5.8), so that the minimal of $\int_{t_{0,p}}^{t_{f,p}} l(\mathbf{x}(t), \mathbf{u}(t)) dt$ can be stably found. Therefore, \mathbf{V}' and $\dot{\mathbf{V}}'$ are constrained in **vehConstraints** considering the maximum and minimum output torque of the powertrain given vehicle speed. The constrained predicted speed profile $\mathbf{V} := [v_{\text{pred},[0]}, v_{\text{pred},[1]}, \dots, v_{\text{pred},[N_p]}]$ and acceleration $\dot{\mathbf{V}} := [\dot{v}_{\text{pred},[0]}, \dot{v}_{\text{pred},[1]}, \dots, \dot{v}_{\text{pred},[N_p]}]$ serve as the input for calculation of the requested torque in (4.4). The constrained predicted speed is referred to as predicted speed in what follows.

predictiveEMS

input: $\mathbf{x}_{[0]}, \mathbf{u}_{[0]}, q_{[0]}, \dot{v}_{\text{req}}, N_p$

$$(\mathbf{V}', \dot{\mathbf{V}}') := \text{predMC}(\mathbf{x}_{[0]}, \dot{v}_{\text{req}}, N_p)$$

$$(\mathbf{V}, \dot{\mathbf{V}}) := \text{vehConstraints}(\mathbf{V}', \dot{\mathbf{V}}')$$

$$(\bar{\omega}^*(\cdot), \mathbf{u}^*(\cdot)) := \text{solveHMP}(\mathbf{V}, \dot{\mathbf{V}}, \mathbf{u}_{[0]}, q_{[0]})$$

Return $\bar{\omega}^*(\cdot)$ and $\mathbf{u}^*(\cdot)$

5.2.2 Conditions of the Hybrid Optimal Control Problem

To apply the HMP to the problem (5.5) with the constraints (5.10)-(5.17) specifically, the indexed costates $\lambda_q := (\lambda_{1,q}, \lambda_{2,q})^T$ are introduced to formulate the indexed Hamiltonians

$$\mathcal{H}_q(\mathbf{x}, \mathbf{u}, \lambda_q) = l(\mathbf{x}, \mathbf{u}) + \lambda_q^T \cdot \mathbf{f}_q(\mathbf{x}, \mathbf{u}),$$

$$= \begin{cases} l(\mathbf{x}, \mathbf{u}) + \lambda_q^T \cdot \frac{b(G, v, T_{\text{total}})}{r} \cdot \begin{pmatrix} i_1 \\ i_2(G) \end{pmatrix}, & \text{if } q \in \mathcal{Q} \setminus \{q_{j,c} \mid 3 < j \leq 9\}, \\ l(\mathbf{x}, \mathbf{u}) + \lambda_q^T \cdot \begin{pmatrix} b(G, v, T_{\text{total}}) \\ r \\ \frac{u_2}{\Theta_{\text{ST2,input}}} \end{pmatrix} \cdot i_1, & \text{if } q \in \{q_{j,c} \mid 3 < j \leq 9\}, \end{cases} \quad (5.18)$$

where the gear position G is determined by the discrete states and sub-states by (4.17), (4.26), (4.31), (4.36) and (4.49). Function b from (4.6) calculates the vehicle acceleration. The backlash means exclude. The second condition is activated, only when the sub-state $q_{j,c}, 3 < j \leq 9$, is activated (see Section 4.4.2.3). The dynamics of the indexed costates that defined in HMP (2.53) are

$$\dot{\lambda}_q(t) = -\frac{\partial \mathcal{H}_q}{\partial \mathbf{x}_q}(\mathbf{x}^*, \mathbf{u}^*, \lambda_q), \quad (5.19)$$

$$\dot{\lambda}_{1,q} = -T_1 - \frac{\partial PL_{EM1}}{\partial \omega_1}(\omega_1, u_1) - \lambda_{1,q} \frac{\partial b}{\partial \omega_1}(G, v, T_{\text{total}}), \quad (5.20)$$

$$\dot{\lambda}_{2,q} = \begin{cases} -T_2 - \frac{\partial PL_{EM2}}{\partial \omega_2}(\omega_2, u_2) - \lambda_{2,q} \frac{\partial b}{\partial \omega_2}(G, v, T_{\text{total}}), & \text{if } q \in \mathcal{Q} \setminus \{q_{j,c} \mid 3 < j \leq 9\}, \\ -T_2 - \frac{\partial PL_{EM2}}{\partial \omega_2}(\omega_2, u_2), & \text{if } q \in \{q_{j,c} \mid 3 < j \leq 9\}. \end{cases} \quad (5.21)$$

$\frac{\partial b}{\partial \omega_1}$ and $\frac{\partial b}{\partial \omega_2}$ are fully expressed in (C.6) in Appendix C. Use an extended state $\mathbf{y}_q := (\mathbf{x}^T, \boldsymbol{\lambda}_q^T)^T$ to keep the notation compact, whose dynamics follow

$$\dot{\mathbf{y}}_q = \boldsymbol{\Gamma}_q(\mathbf{y}_q, \mathbf{u}_q^*) = \begin{pmatrix} \mathbf{f}_q(\mathbf{x}, \mathbf{u}_q^*) \\ -\frac{\partial \mathcal{H}_q}{\partial \mathbf{x}_q}(\mathbf{x}^*, \mathbf{u}_q^*, \boldsymbol{\lambda}_q) \end{pmatrix}. \quad (5.22)$$

The hybrid controls $(\varpi^*(t), \mathbf{u}^*(t))$ from (2.55) are determined by firstly

$$\mathbf{u}_q^* = \underset{\mathbf{u}_q \in \mathcal{U}}{\operatorname{argmin}} \mathcal{H}_q(\mathbf{x}, \mathbf{u}_q, \boldsymbol{\lambda}_q), \quad (5.23)$$

subject to the constraints (5.13) and (5.14). Secondly, the discrete control is determined by

$$\varpi^* = \underset{\varpi \in \Pi_q}{\operatorname{argmin}} \mathcal{H}_q(\mathbf{x}, \mathbf{u}_q^*, \boldsymbol{\lambda}_q), \quad (5.24)$$

$$\mathbf{u}^* = \mathbf{u}_{\varpi^*}^*. \quad (5.25)$$

The transversality condition (2.55) of the problem (5.5) is

$$\boldsymbol{\lambda}_{q(t_f)}(t_f) - 2\beta(\mathbf{x}(t_{f,p}) - \mathbf{x}_{t_{f,p}}) = \mathbf{0}, \quad (5.26)$$

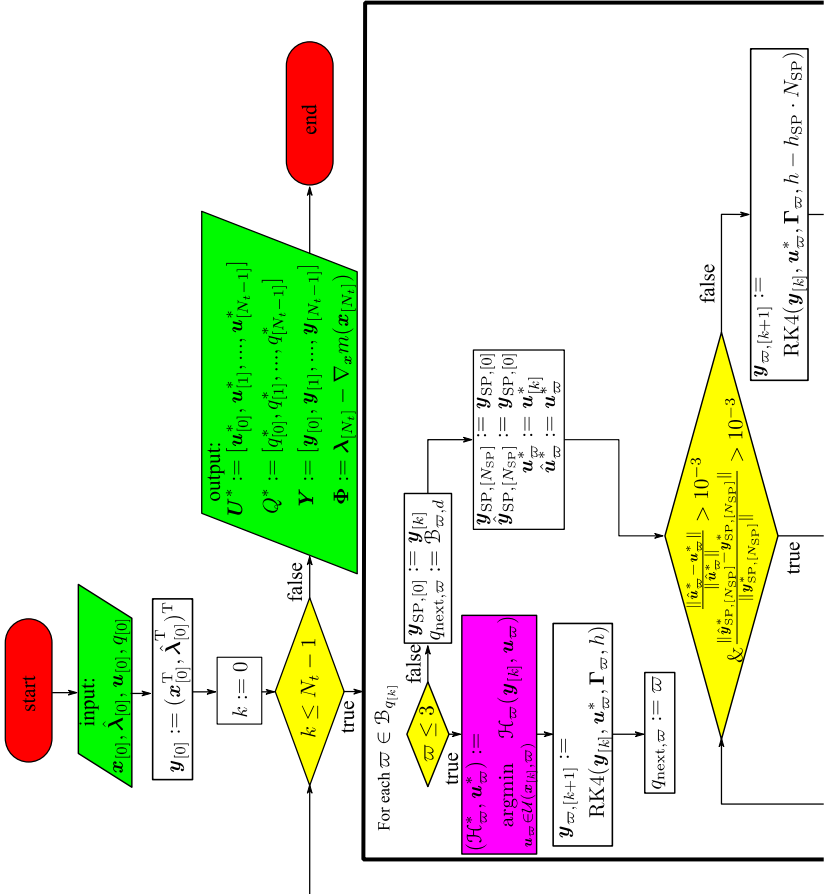
which is evaluated by its error

$$\boldsymbol{\Phi} = \boldsymbol{\lambda}_{q(t_f)}(t_f) - \nabla_{\mathbf{x}} m(\mathbf{x}^*(t_f)). \quad (5.27)$$

5.2.3 Algorithm for Solving the Hybrid Optimal Control Problem

The function **shootHMP** determines a candidate solution of HMP, which produces a pair of continuous control function and discrete state sequence on the discrete time (U^*, Q^*) , given a presumed initial costate $\hat{\lambda}_{[0]}$ and the initial state, discrete state and control $(x_{[0]}, q_{[0]}$ and $u_{[0]})$. The dynamics from (5.22) are calculated by explicit fourth-order Runge-Kutta integration scheme (RK4) described in Section 2.3. Its algorithm is omitted. The candidate is considered a solution, when the norm $\|\Phi\|$ is smaller than the convergence limit ε , which is evaluated in the function **solveHMP**.

Figure 5.3 shows the flowchart of **shootHMP**. The timestep size of the discrete state sequence (h) is 1 s. At every timestep, ϖ is enumerated in the set $\mathcal{B}_{q_{[k]}}$ listed in Table 4.2. When the power-train stays in the same gear, the function “argmin” returns both the minimum value of the indexed



Hamiltonian \mathcal{H}_{ϖ}^* and the corresponding control \mathbf{u}_{ϖ}^* (marked magenta), which is used to calculate the dynamics for the indexed extended states at next timestep $\mathbf{y}_{\varpi,[k+1]}$. During a shift process, i.e. $\varpi > 3$, the state trajectory is calculated with a time step (h_{SP}) of 0.01 s. The controls are predefined by the shift process in 4.4.2 and the dynamics of the state are therefore autonomous. N_{SP} is the time index, when the autonomous switching manifold (4.50) is met, i.e. a shift process finishes. q_{next} is the gear position at which the shift process ends. It is defined by $\mathcal{B}_{\varpi,d}$ in Table 4.3. \mathbf{u}_{ϖ}^* is determined by minimizing the Hamiltonian in the target gear, which gives the $T_{1,\text{new}}$ and $T_{2,\text{new}}$ in Subsection 4.4.2.4. A “while” loop is performed until the control (\mathbf{u}_{ϖ}^*) and the state ($\mathbf{y}_{\text{SP},[N_{\text{SP}]}$) converge, which is evaluated by the relative change of their Euclidean norm ($\|\cdot\|$). The dynamics are calculated in the rest of the time ($h - h_{\text{SP}} \cdot N_{\text{SP}}$) to determine $\mathbf{y}_{\varpi,[k+1]}$. \mathcal{H}_{ϖ}^* is approximated by the values during the shift process and the Hamiltonian in the q_{next} together (marked magenta).

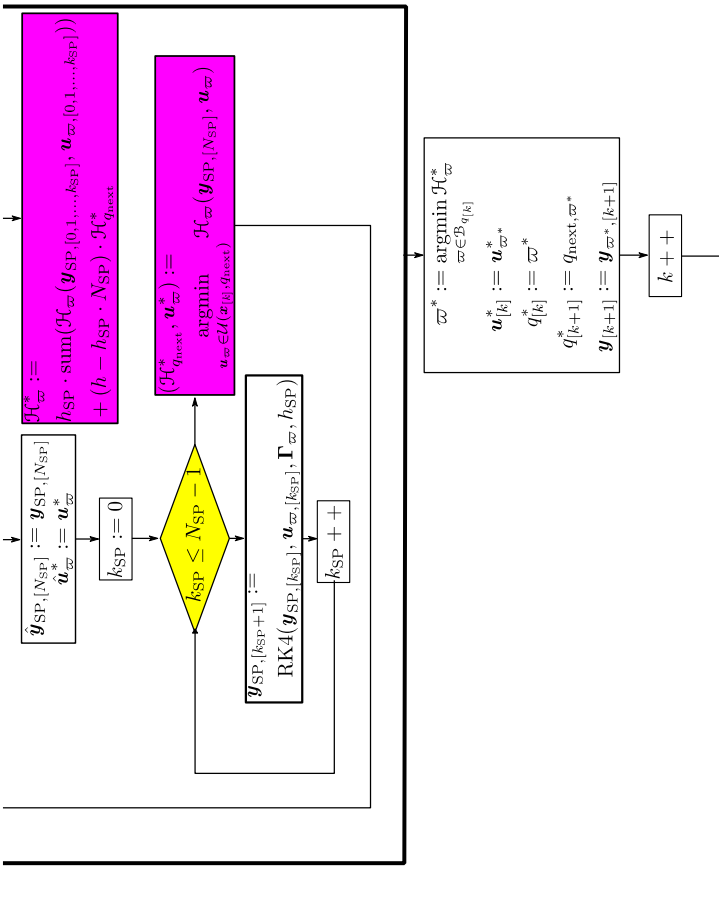


Figure 5.3: Flowchart of shootHMP

After the enumeration of ϖ , ϖ^* is chosen by finding the smallest indexed Hamiltonian. $\mathbf{u}_{[k]}^*$, $q_{[k]}^*$, $q_{[k+1]}^*$ and $\mathbf{y}_{[k+1]}$ are determined accordingly.

Broyden method [143], a quasi-Newton method, is used to find the root of Φ . The initial costates are updated iteratively according to

$$\hat{\lambda}_{[0],n+1} = \hat{\lambda}_{[0],n} - J_n^+ \cdot \Phi_n, \quad (5.28)$$

$$J_n^+ = J_{n-1}^+ + \frac{\Delta \hat{\lambda}_{[0],n} - J_{n-1}^+ \cdot \Delta \Phi_n}{\Delta \hat{\lambda}_{[0],n}^T \cdot J_{n-1}^+ \cdot \Delta \Phi_n} \Delta \hat{\lambda}_{[0],n}^T \cdot J_{n-1}^+, \quad (5.29)$$

$$\Delta \hat{\lambda}_{[0],n} = \hat{\lambda}_{[0],n} - \hat{\lambda}_{[0],n-1}, \quad (5.30)$$

$$\Delta \Phi_n = \Phi_n - \Phi_{n-1}, \quad (5.31)$$

where n is the iteration index. J_n^+ , the pseudoinverse [144] of the finite-difference approximated Jacobian matrix J_n , can be iteratively obtained without updating the Jacobian matrix. J is initialized with a 2×2 identity matrix I_2 . See [143] for detailed explanation. The pseudoinverse is used to replace the inverse in the original algorithm, so that the Broyden method can be executed, even when the matrix J_n is non invertible. The Broyden method is implemented in **solveHMP**, whose flowchart is shown in Figure 5.4, where **broydenJacobian** calculates J_n^+ based on (5.29)-(5.31). The convergence criterion depends on the initial guesses $\hat{\lambda}_{[0]}$, which are generated by the function **genInitialCostate** described in Subsection 5.2.4.

5.2.4 Initialization Inspired by Bang-Bang Controls

The guesses are initialized based on bang-bang controls, i.e. the control switches between upper and lower bounds [145]. For the problem in question, it can be interpreted as that either EM provides the total output torque on the wheels alone.

Figure 5.5 shows the flowchart of the function **genInitialCostate**. The Broyden method requires two initial guesses, which eliminates q_3 , with which only one candidate with EM1 propelling the vehicle is available. If the system starts with q_3 , the discrete state q that is used in **genInitialCostate** is either q_1 or q_2 , chosen by the smaller cost function at the initial time point. The state trajectory over the prediction horizon $\mathbf{X} := [\mathbf{x}_{[0]}, \mathbf{x}_{[1]}, \dots, \mathbf{x}_{[N_t]}]$ is determined by the predicted speed without calculating the dynamics. The output torque on the wheels over the prediction horizon $\mathbf{T}_{\text{total}}$, similarly defined as \mathbf{X} , is determined by (4.4)-(4.6). An example of \mathbf{X} and $\mathbf{T}_{\text{total}}$ are shown in Figure 5.6(a). Based on it, the torques of the EM1 and the EM2 providing the total output torque on the wheels alone, $\mathbf{T}_{1\text{alone}}$ and $\mathbf{T}_{2\text{alone}}$ are calculated and are used to form the control functions $\mathbf{U}_{1\text{alone}}$ and $\mathbf{U}_{2\text{alone}}$, respectively. The power losses in STs are neglected. Should the total output torque on the wheels be not reached in either case, it is complemented by the other EM.

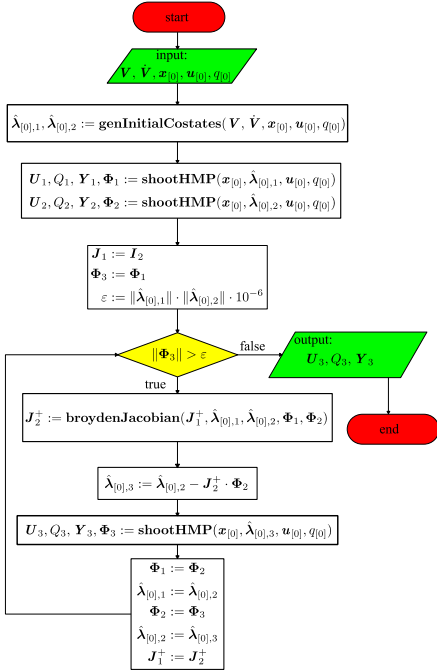


Figure 5.4: Flowchart of solveHMP

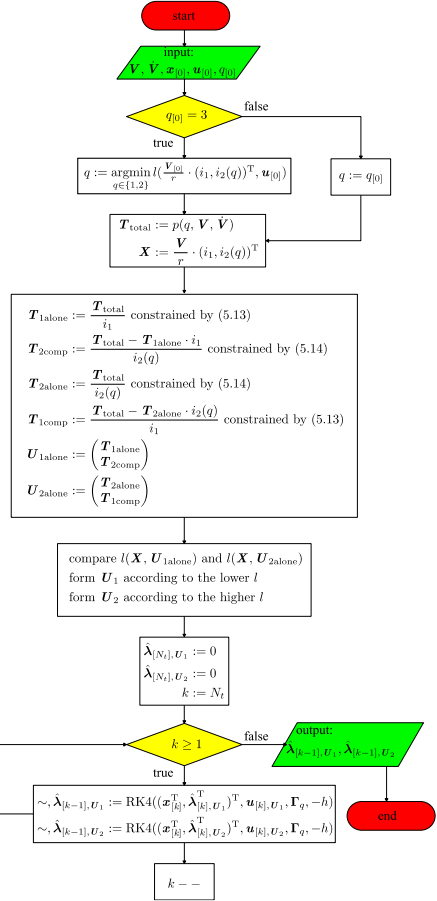


Figure 5.5: Flowchart of genInitialCostate

The control function for the first guess \mathbf{U}_1 is formed by choosing the controls from $\mathbf{U}_{1\text{alone}}$ and $\mathbf{U}_{2\text{alone}}$ when they cause a smaller cost defined by (5.7). $\tau_{1\text{alone}}$ denotes the set of the time indexes when the condition in “find” is fulfilled. $[0, 1, \dots, N_t] / \tau_{1\text{alone}}$ denotes the set of the time indexes that are not included in $\tau_{1\text{alone}}$. Switch the sets and form the control function \mathbf{U}_2 for the second guess. Figure 5.6(b) shows an example of $\mathbf{U}_{1\text{alone}}$, $\mathbf{U}_{2\text{alone}}$ and the deviation of cost function along the prediction horizon, $l(\mathbf{X}, \mathbf{U}_{1\text{alone}}) - l(\mathbf{X}, \mathbf{U}_{2\text{alone}})$. \mathbf{U}_1 is $\mathbf{U}_{1\text{alone}}$, when the deviation is smaller than or equal to zero, and $\mathbf{U}_{2\text{alone}}$, when the deviation is positive. \mathbf{U}_2 is the opposite order. An example is shown in Figure 5.6(c). The initial values of costates are calculated backwards by (5.19) with explicit Euler method, whose example is shown in Figure 5.6(d).

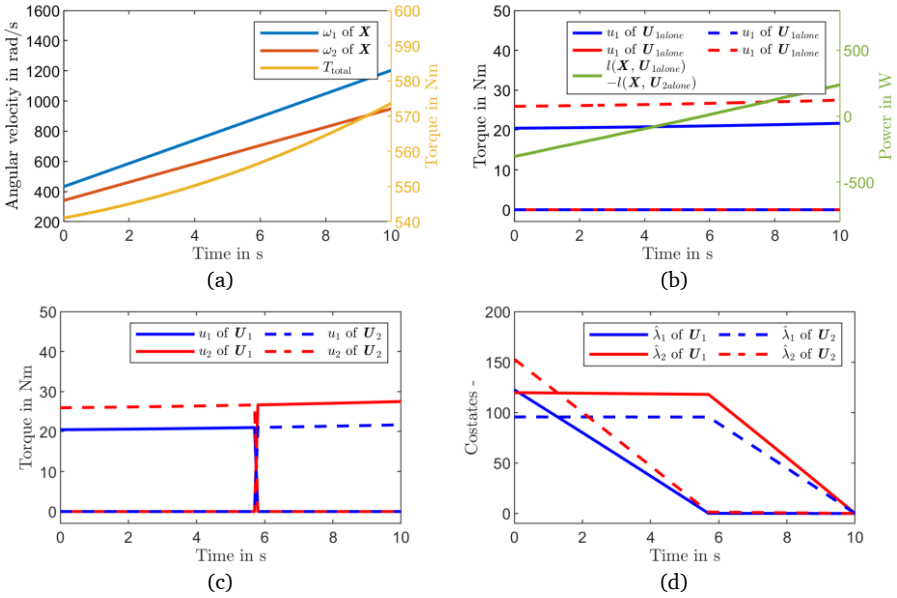


Figure 5.6: An example of the initialization for a predicted speed profile with constant acceleration and the system is in q_2 . (a): predicted X and T_{total} . (b): approximated U_{1alone} and U_{2alone} . Compare the cost function over the prediction horizon. (c): form U_1 and U_2 . (d) estimate $\hat{\lambda}_{i_1, U_1}$, $\hat{\lambda}_{i_1, U_2}$ by backwards integration.

5.2.5 Minimization of the Hamiltonian

Given (5.18) and a determined q , the minimization of the Hamiltonian (5.23) is expressed as

$$\begin{aligned}
 \min_{\mathbf{u}_q} \mathcal{H}_q(\mathbf{x}, \mathbf{u}_q, \boldsymbol{\lambda}_q) &= \min_{\mathbf{u}_q} \left(l(\mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}_q^T \cdot \frac{b(G, v, p(G, v, \dot{v}_{pred}))}{r} \cdot \begin{pmatrix} i_1 \\ i_2(q) \end{pmatrix} \right) \\
 &= \min_{\mathbf{u}_q} \left(l(\mathbf{x}, \mathbf{u}_q) + \boldsymbol{\lambda}_q^T \cdot \frac{b(G, v, p(G, v, \dot{v}_{pred}))}{r} \cdot \begin{pmatrix} i_1 \\ i_2(q) \end{pmatrix} \right), \quad (5.32)
 \end{aligned}$$

which tells that the task is to minimize the cost function (5.7) over the control \mathbf{u}_q (torques of EM1 and EM2) for a given discrete state q .

In Section 5.3, the minimization of the cost function is studied.

5.3 Minimization of the Cost Function

The minimization of the cost function is formulated as

$$\min_{\mathbf{u}_q} l(\mathbf{x}, \mathbf{u}_q), \quad (5.33)$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + PL_{EM1}(\omega_1, u_1) + PL_{EM2}(\omega_2, u_2), \quad (5.7)$$

subject to

$$u_1 \in [T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)], \quad (5.8)$$

$$u_2 \in [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)], \quad (5.9)$$

$$\dot{\mathbf{x}}_q = \mathbf{f}_q(\mathbf{x}, T_{\text{total}}) \quad (4.14)$$

$$= \frac{b(q, v, T_{\text{total}})}{r} \cdot \begin{pmatrix} i_1 \\ i_2(q) \end{pmatrix}, \quad (4.15)$$

$$T_{\text{total}} = \phi_1(\omega_1, u_1) \cdot i_1 + \phi_{2,q}(\omega_2, u_2) \cdot i_2(q) \quad (4.16)$$

$$q \in \{1,2\}, \quad (*)$$

where in (*) only q_1 and q_2 are considered, since the torques are determined by (4.7)-(4.9) in neutral gear position and the continuous controls are predefined during a shift process ($q \in \mathcal{Q}\{q_1, q_2, q_3\}$). The minimization is subject to the inequality constraints (5.8) and (5.9) as well as the equality constraint (4.16). (5.15) and (5.16) are not considered, since \mathbf{x} and q are given. Due to the nonlinearity of the power losses functions in the cost function, the best-case scenario is to have the problem solved with efficient algorithms for convex optimization problems (see discussion in Section 5.5). Convex optimization problems refer to optimizing convex functions on convex sets [63, pp. 136-137]. The cost function and the constraints of the problem (5.33) is shown in Figure 5.7(a). The axes are the torques of both EM. The black square sketches the contour of the

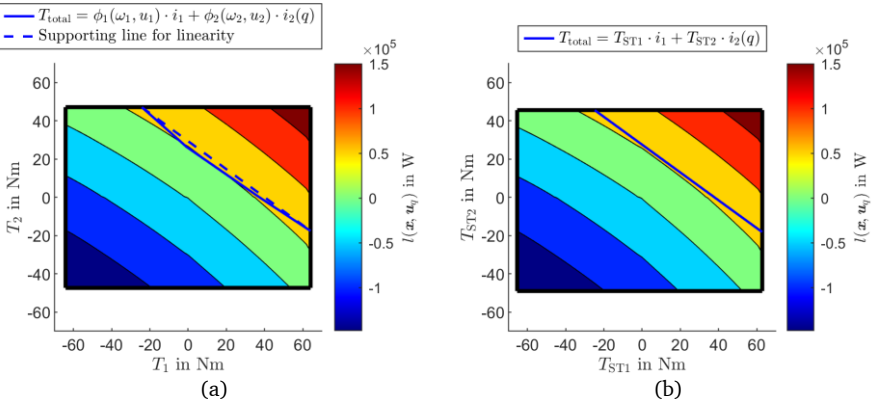


Figure 5.7: The constraints of the minimization problem with following parameters: $v = 60$ km/h, $T_{\text{total}} = 1000$ Nm, $q = q_1$. (a) problem (5.33); (b) equivalent problem.

cartesian product $[T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)] \times [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)]$ from (5.8) and (5.9), which is a convex set. The blue solid curve represents the equality constraint (4.16), while the dotted line represents a line segment connecting both ends of the curve. Evidently, the optimization domain that is constructed by (5.8), (5.9) and (4.16) is not a convex set (see its definition in [63, pp. 21-26]). The problem (5.33) is not a convex optimization problem. An NLP problem needs to be solved at each time step in **shootHMP**.

The minimization (5.33) can be proved to be equivalent to a convex optimization problem in standard form, i.e. minimizing a convex function subject to convex inequality constraints and affine equality constraints [63, pp. 136-137], which makes the use of efficient and broadly available optimization algorithms, e.g. newton's method [63, pp. 525-531] and interior-point method [146, pp. 412-415] possible. Furthermore, additional inequality constraints can be found to make the optimization domain smaller. The proofs are elaborated in Appendix D. The following subsections provide the results of the proofs and the main idea of the proofs for better understanding.

5.3.1 Equivalent Convex Optimization

Lemma 1: the problem (5.33) is equivalent to a convex optimization problem in a standard form.

Proof: See Proof of Lemma 1 in Appendix D.

Main idea of the proof: change the variables to the effective torques of both STs, i.e.

$$\min_{(T_{ST1}, T_{ST2})^T} l_{x,q}(\phi_{1,\omega_1}^{-1}(T_{ST1}), \phi_{2,\omega_2,q}^{-1}(T_{ST2})), \quad (5.34)$$

$$\begin{aligned} l_{x,q}(\phi_{1,\omega_1}^{-1}(T_{ST1}), \phi_{2,\omega_2,q}^{-1}(T_{ST2})) \\ = \omega_1 \cdot \phi_{1,\omega_1}^{-1}(T_{ST1}) + \omega_2 \cdot \phi_{2,\omega_2,q}^{-1}(T_{ST2}) \\ + PL_{EM1}(\omega_1, \phi_{1,\omega_1}^{-1}(T_{ST1})) + PL_{EM2}(\omega_2, \phi_{2,\omega_2,q}^{-1}(T_{ST2})), \end{aligned} \quad (5.35)$$

subject to

$$T_{ST1} \in \{\phi_{ST1,\omega_1}(T) \mid T \in \mathcal{T}_{EM1,\omega_1}\}, \quad (5.36)$$

$$T_{ST2} \in \{\phi_{ST2,\omega_2}(T) \mid T \in \mathcal{T}_{EM2,\omega_2}\}, \quad (5.37)$$

$$\dot{x}_q = f_q(x, T_{total}) \quad (5.38)$$

$$= \frac{b(q, v, T_{total})}{r} \cdot \begin{pmatrix} i_1 \\ i_2(q) \end{pmatrix}, \quad (5.39)$$

$$T_{total} = T_{ST1} \cdot i_1 + T_{ST2} \cdot i_2(q), \quad (4.7)$$

$$q \in \{1, 2\}, \quad (*)$$

where x and q are subscripted in (5.34) and (5.35) to avoid confusion, since the cost function is optimized w.r.t. torques. The function $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$ in (4.16) is more rigidly defined as

$$T_{ST} = \phi(\omega, T) = T - \frac{PL_{ST}(\omega, T)}{\omega}, T \in \mathcal{T}(\omega), \quad (5.40)$$

$$\mathcal{T}(\omega) = \{T \in \mathbb{R} \mid T_{min}(\omega) \leq T \leq T_{max}(\omega)\}. \quad (5.41)$$

Since the angular velocity ω is given, the function is rewritten as $\phi_\omega: \mathbb{R} \rightarrow \mathbb{R}$

$$T_{ST} = \phi_\omega(T) = T - \frac{PL_{ST}(\omega, T)}{\omega}, T \in \mathcal{T}_\omega, \quad (5.42)$$

$$\mathcal{T}_\omega = \{T \in \mathbb{R} \mid T_{min}(\omega) \leq T \leq T_{max}(\omega)\}, \quad (5.43)$$

where the domain \mathcal{T}_ω is bounded by the maximum and the minimum torque of the EM depending on ω . ϕ_ω^{-1} is the inverse function of ϕ_ω . Its existence and properties are discussed in Appendix D.

Figure 5.7(b) shows the cost function on the axes of the effective torques of both STs. The blue line represents the equality constraint (4.7), which is affine. The optimization domain constructed by (5.36), (5.37) and (4.7) is convex. By definition of convexity and the examination of ϕ_ω^{-1} , the function to be minimized in (5.34) can be proved to be convex. (5.34) is a convex optimization problem in standard form. The equivalency is shown in Appendix D.

5.3.2 Upper and Lower Bounds of the Torques

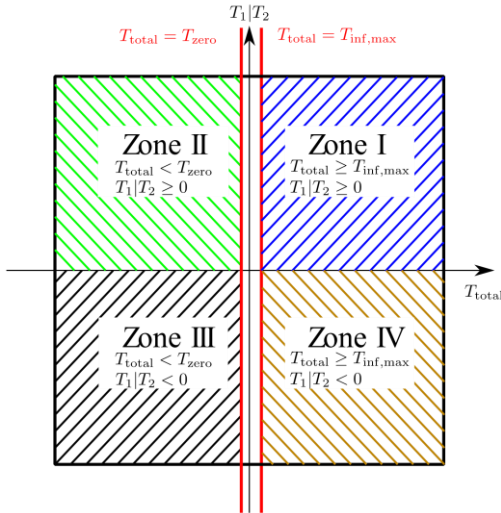


Figure 5.8: Torque range divided into several zones

The domain of the optimization problem (5.33) and (5.34) are studied and reduced, so that the optimization can be faster. According to (5.13) and (5.14), \mathbf{u}_q^* falls into the cartesian product of the space $[T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)] \times [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)]$ with given $\mathbf{x} = (\omega_1, \omega_2)^T$ and q . To reduce the domain is to prove that \mathbf{u}_q^* only falls into a subset of the cartesian product of the space. Furthermore, with the equality constraint (4.16) in mind, T_{total} also needs to be considered. Intuitively, if T_{total} is a large positive real number, both u_1 and u_2 are likely to be positive, and vice versa. Therefore, instead of $(u_1, u_2)^T$, two vectors $(T_{\text{total}}, u_1)^T$ and $(T_{\text{total}}, u_2)^T$ are examined.

For a given \mathbf{x} and $q < 3$, the cartesian product of the space $[T_{\text{total},\min}(\omega_1, \omega_2), T_{\text{total},\max}(\omega_1, \omega_2)] \times [T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)]$ is marked with a black square in Figure 5.8 with T_{total} as x-axis. For simple display, T_1 and T_2 share y-axis. $[T_{\text{total},\min}(\omega_1, \omega_2), T_{\text{total},\max}(\omega_1, \omega_2)] \times [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)]$ with a black square. By equality constraint (4.16), there is

$$T_{\text{total},\min}(\omega_1, \omega_2) = \phi_1(\omega_1, T_{1,\min}(\omega_1)) \cdot i_1 + \phi_{2,q}(\omega_2, T_{2,\min}(\omega_2)) \cdot i_2(q), \quad (5.44)$$

$$T_{\text{total},\max}(\omega_1, \omega_2) = \phi_1(\omega_1, T_{1,\max}(\omega_1)) \cdot i_1 + \phi_{2,q}(\omega_2, T_{2,\min}(\omega_2)) \cdot i_2(q). \quad (5.45)$$

The torque of an EM minimizing the power loss of an EM or a ST individually is expressed as

$$T_{1,\text{EM1},\text{inf}} = \underset{T_1}{\operatorname{argmin}} PL_{\text{EM1}}(\omega_1, T_1), \quad (5.46)$$

$$T_{2,\text{EM2},\text{inf}} = \underset{T_2}{\operatorname{argmin}} PL_{\text{EM2}}(\omega_2, T_2), \quad (5.47)$$

$$T_{1,\text{ST1},\text{inf}} = \underset{T_1}{\operatorname{argmin}} PL_{\text{ST1}}(\omega_1, T_1), \quad (5.48)$$

$$T_{2,\text{ST2},\text{inf}} = \underset{T_2}{\operatorname{argmin}} PL_{\text{ST2},q}(\omega_2, T_2), \quad (5.49)$$

$$T_{1,\text{inf}} = \max(T_{1,\text{EM1},\text{inf}}, T_{1,\text{ST1},\text{inf}}) = T_{1,\text{EM1},\text{inf}}, \quad (5.50)$$

$$T_{2,\text{inf}} = \max(T_{2,\text{EM2},\text{inf}}, T_{2,\text{ST2},\text{inf}}) = T_{2,\text{EM2},\text{inf}}, \quad (5.51)$$

where the subscript ‘‘inf’’ denote infimum. For instance, $T_{1,\text{EM1},\text{inf}}$ denotes the torque of EM1 that minimizes the power loss of EM1. Examine the power losses of both STs and both EMs (see Figure D. 1), $0 \leq T_{1,\text{ST1},\text{inf}} < T_{1,\text{EM1},\text{inf}}$ and $0 \leq T_{2,\text{ST2},\text{inf}} < T_{2,\text{EM2},\text{inf}}$, which gives (5.50) and (5.51). The output torques that is solely provided by either of the EM with the torques in (5.50) and (5.51) are

$$T_{\text{total},1,\text{inf}} = \phi_{\omega_1,\text{ST1}}(T_{1,\text{inf}}) \cdot i_1 + \phi_{\omega_2,\text{ST2},q}(0) \cdot i_2(q), \quad (5.52)$$

$$T_{\text{total},2,\text{inf}} = \phi_{\omega_1,\text{ST1}}(0) \cdot i_1 + \phi_{\omega_2,\text{ST2},q}(T_{2,\text{inf}}) \cdot i_2(q), \quad (5.53)$$

$$T_{\text{total},\text{inf},\text{max}} = \max(T_{\text{total},1,\text{inf}}, T_{\text{total},2,\text{inf}}), \quad (5.54)$$

where $T_{\text{total},\text{inf},\text{max}}$ is the bigger one of the two output torques. Additionally, define the output

torque on the wheels that both EM output zero torque according to (4.16) as

$$T_{\text{zero}} = \phi_{\omega_1, ST1}(0) \cdot i_1 + \phi_{\omega_2, ST2, q}(0) \cdot i_2(q). \quad (5.55)$$

The lines $T_{\text{total}} = T_{\text{inf,max}}$ and $T_{\text{total}} = T_{\text{zero}}$ are marked red in Figure 5.8. The cartesian product of the spaces are then divided into several zones. It is firstly proved for Lemma 2, that $(T_{\text{total}}, u_1^*)$ and $(T_{\text{total}}, u_2^*)$ fall into certain Zones, when $T_{\text{total}} < T_{\text{zero}}$ or $T_{\text{total}} \geq T_{\text{total,inf,max}}$, which gives additional constraints for the problems (5.33) and (5.34). $T_{\text{zero}} \leq T_{\text{total}} < T_{\text{total,inf,max}}$ is later addressed in Lemma 3, since the domain of the optimization problem in this zone is different from the rest.

Lemma 2: Suppose $(T_{ST1}^*, T_{ST2}^*)^T$ solves (5.34) and $(u_1^*, u_2^*)^T = (\phi_{ST1, \omega_1}^{-1}(T_{ST1}^*), \phi_{ST2, \omega_2}^{-1}(T_{ST2}^*))^T$ is the equivalent solution to (5.33). Following inequality holds

$$\begin{cases} u_1^* \in [0, T_{1,\text{max}}(\omega_1)], \\ u_2^* \in [0, T_{2,\text{max}}(\omega_2)], \\ u_1^* \in [T_{1,\text{min}}(\omega_1), 0], \\ u_2^* \in [T_{2,\text{max}}(\omega_2), 0], \end{cases} \quad T_{\text{total}} \geq T_{\text{total,inf,max}}, \quad (5.56)$$

$$\begin{cases} T_{ST1}^* \in [\phi_{\omega_1, ST1}(0), \phi_{\omega_1, ST1}(T_{1,\text{max}}(\omega_1))], \\ T_{ST2}^* \in [\phi_{\omega_2, ST2, q}(0), \phi_{\omega_2, ST2, q}(T_{2,\text{max}}(\omega_2))], \\ T_{ST1}^* \in [\phi_{\omega_1, ST1}(T_{1,\text{min}}(\omega_1)), \phi_{\omega_1, ST1}(0)], \\ T_{ST2}^* \in [\phi_{\omega_2, ST2, q}(T_{2,\text{max}}(\omega_2)), \phi_{\omega_2, ST2, q}(0)], \end{cases} \quad T_{\text{total}} < T_{\text{zero}}. \quad (5.57)$$

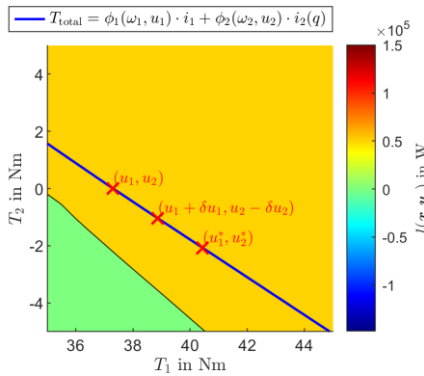


Figure 5.9: Illustration for Proof of Lemma 2.

Proof: See Proof of Lemma 2 in Appendix D.

Main idea of the proof: Consider $T_{\text{total}} \geq T_{\text{inf,max}}$, at least one of $(T_{\text{total}}, u_1^*)$ and $(T_{\text{total}}, u_2^*)$ fall in “Zone I”. Due to the equality constraint (4.16), they cannot fall in “Zone IV” together. It can be proved by contradiction that they fall in “Zone I”.

Assume $u_1^* \geq 0$, i.e. “Zone I”, and $u_2^* < 0$, i.e. “Zone IV”. Choose a $u_1 \in [0, u_1^*]$ in such a way that $u_2 = 0$ can be determined by (4.16). A point $(u_1 + \delta u_1, u_2 - \delta u_2)^T, \delta u_1 > 0, \delta u_2 > 0^{48}$, fulfils (4.16) and locates between $(u_1, u_2)^T$ and $(u_1^*, u_2^*)^T$. These three points are marked in Figure 5.9, an enlarged view of Figure 5.7(a) without the supporting line.

$\phi_{\omega_1, \text{ST1}}$ and $\phi_{\omega_2, \text{ST2}, q}$ map $(u_1, u_2)^T, (u_1 + \delta u_1, u_2 - \delta u_2)^T$ and $(u_1^*, u_2^*)^T$ to $(T_{\text{ST1}}, T_{\text{ST2}})^T, (T_{\text{ST1}} + \delta T_{\text{ST1}}, T_{\text{ST2}} - \delta T_{\text{ST2}})^T$ and $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$, respectively, where $\delta T_{\text{ST1}} > 0$ and $\delta T_{\text{ST2}} > 0$. They fulfil the affine equality constraint (4.7). Since (5.34) is a convex function,

$$l_{x,q} \left(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}}), \phi_{2,\omega_2,q}^{-1}(T_{\text{ST2}}) \right) \geq l_{x,q} \left(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}} + \delta T_{\text{ST1}}), \phi_{2,\omega_2,q}^{-1}(T_{\text{ST2}} - \delta T_{\text{ST2}}) \right). \quad (5.58)$$

However, (5.58) is shown to be false in Proof of Lemma 2. Ergo, the assumption does not hold.

The above process is repeated with modification for the case of $u_1^* < 0$, i.e. “Zone IV”, and $u_2^* \geq 0$, i.e. “Zone I”.

Therefore, $(T_{\text{total}}, u_1^*)$ and $(T_{\text{total}}, u_2^*)$ cannot fall in “Zone I” and “Zone IV” separately. When $T_{\text{total}} \geq T_{\text{inf,max}}$, both $(T_{\text{total}}, u_1^*)$ and $(T_{\text{total}}, u_2^*)$ fall in “Zone I”.

The same is done for the case that $T_{\text{total}} < T_{\text{zero}}$.

For the range that $T_{\text{zero}} \leq T_{\text{total}} < T_{\text{total,inf,max}}$, the following constraints can be found:

Lemma 3: Suppose $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$ solves (5.34) and $(u_1^*, u_2^*)^T = \left(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}}^*), \phi_{2,\omega_2}^{-1}(T_{\text{ST2}}^*) \right)^T$ is the equivalent solution to (5.33). Following inequality holds

$$\begin{aligned} u_1^* &\in \left[\phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1} \right), T_{1,\text{inf,max}} \right], \\ u_2^* &\in \left[\phi_{2,\omega_2,q}^{-1} \left(\frac{T_{\text{total}} - \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1}{i_2(q)} \right), T_{2,\text{inf,max}} \right], \end{aligned} \quad T_{\text{zero}} \leq T_{\text{total}} < T_{\text{total,inf,max}}, \quad (5.59)$$

⁴⁸ See Proof of Lemma 2 for rigid definition.

$$\begin{aligned}
T_{ST1}^* &\in \left[\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1}, \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \right], \\
T_{ST2}^* &\in \left[\frac{T_{\text{total}} - \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1}{i_2(q)}, \phi_{2,\omega_2,q}^{-1}(T_{2,\text{inf,max}}) \right], \quad T_{\text{zero}} \leq T_{\text{total}} < T_{\text{total,inf,max}}. \quad (5.60)
\end{aligned}$$

$T_{1,\text{inf,max}}$ and $T_{2,\text{inf,max}}$ denote the torque of the EM1 and the EM2 providing a total output torque of $T_{\text{total,inf,max}}$, which are

$$T_{1,\text{inf,max}} = \phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total,inf,max}} - \phi_{2,\omega_2,q}(0) \cdot i_2(q)}{i_1} \right), \quad (5.61)$$

$$T_{2,\text{inf,max}} = \phi_{2,\omega_2,q}^{-1} \left(\frac{T_{\text{total,inf,max}} - \phi_{1,\omega_1}(0) \cdot i_1}{i_2(q)} \right). \quad (5.62)$$

Proof: See Proof of Lemma 3 in Appendix D.

Main idea of the proof: Lemma 3 can be proved by contradiction.

Assume $u_1^* > T_{1,\text{inf,max}}$. By doing so, (4.16) determines $u_2^* < 0$ (torque of the EM2). Choose a small increment $\delta u_1 > 0$ in such a way that $u_1^* - \delta u_1 > T_{1,\text{inf,max}}$. The equality constraint (4.16) determines the new torque of EM2 corresponding to $u_1^* - \delta u_1$ and is denoted as $u_2^* + \delta u_2$, $\delta u_2 > 0$, which satisfies $u_2^* + \delta u_2 < 0$. By optimality,

$$l_{x,q}((u_1 - \delta u_1, u_2 + \delta u_2)^T) \geq l_{x,q}((u_1^*, u_2^*)^T), \quad (5.63)$$

which can be shown false in Proof of Lemma 3. Therefore, $u_1^* \leq T_{1,\text{inf,max}}$.

Similarly, it can be shown that $u_1^* \geq \phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total}} - \phi_{\omega_2,ST2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1} \right)$.

5.4 Time Complexity of the Predictive EMS

The variables that influence the time complexity of the algorithm `solveHMP` are listed in Table 5.1. As discussed in Section 5.2.3, N_{SP} is dependent on the shift durations. Since the worst-case time complexity, i.e. big-O notation⁴⁹, is discussed, it is assumed that a shift process takes up a time step h . An integration scheme of N_{RK} order is applied to calculate the dynamics.

⁴⁹ To state that the time complexity considering n as its variable is $O(g(n))$ means that there exists a constant C and $C \cdot g(n)$ is an upper bound of the time complexity [81, pp. 34-37].

Table 5.1: Variables influencing the time complexity of solveHMP

| Variable | Description |
|---|--|
| N_{TPBVP} | Number of the iterations to solve the TPBVP |
| N_p | Length of the prediction horizon |
| N_B | Size of the admissible discrete control set |
| $N_{\text{SP}} = \frac{h}{h_{\text{SP}}}$ | Number of the time steps with a length of h_{SP} that a shift process takes |
| $N_{\text{RK}} = 4$ | Order of the RK integration scheme |
| N_{G_t} | Number of the intervals of the time discretization grid for a time step in the RK integration scheme |
| N_{while} | Number of the iterations for the “while” loop in Figure 5.3 to converge |
| N_u | Dimension of the continuous controls |
| N_x | Dimension of the continuous states |
| N_y | Dimension of the extended states consisting of \mathbf{x} and λ |

The time complexity of **solveHMP** is the summary of five parts that are listed in Table 5.2. $\text{time}_{\text{min}l}$ and $\text{time}_{\text{min}\mathcal{H}}$ are the time complexity of the minimization of the cost function and of the Hamiltonian, respectively. They will be elaborated in the next paragraph. It is assumed that the time complexity of all mathematical operations and logic operations cost the same complexity. In **shootHMP**, the admissible discrete control set is enumerated. (5.65) estimates the time complexity of determining the trajectories of \mathbf{x} and λ as well as \mathbf{u}_q^* through $\text{min}\mathcal{H}$ for $\varpi \leq 3$, i.e. the gear position does not change. (5.66) estimates the time complexity of the same procedure but for $\varpi > 3$, i.e. the gear position changes. After each time step, “Evaluate” takes places. It finds the smallest \mathcal{H} out of a vector of N_B elements.

$N_{\text{RK}} \cdot N_y \cdot N_{G_t} \cdot C_{\text{RK}}$ calculates the complexity of the RK scheme. C denotes a constant, which is used to summarize the number of the computation steps in a process that are independent on the variables in Table 5.1. The subscripts of C indicate the process within which the operations are summarized. For instance, C_{RK} is the constant for the RK integration scheme, whose value is estimated to be 36 for calculating \mathbf{x} and λ at each step in the RK integration scheme. The time complexity of searching in a look-up table grows linearly with the number of the points in its discretization grid, which is included in $C_{\varpi \leq 3}$, $C_{\varpi > 3}$ and C_{Eva} .

Table 5.2: Time complexity of each part in solveHMP

| | | | |
|-----------------|-------------------|---|--------|
| Initialization: | | $\text{time}_{\min l} + N_{\mathbf{u}} \cdot N_p \cdot C_{\text{ini}} + N_p \cdot (N_{\text{RK}} \cdot N_y \cdot N_{g_t} \cdot C_{\text{RK}})$ | (5.64) |
| shootHMP | $\varpi \leq 3$: | $N_{\text{TPBVP}} \cdot N_p \cdot [\text{time}_{\min \mathcal{H}} + (N_{\text{RK}} \cdot N_y \cdot N_{g_t} \cdot C_{\text{RK}}) + C_{\varpi \leq 3}]$ | (5.65) |
| | $\varpi > 3$: | $N_{\text{TPBVP}} \cdot N_p \cdot (N_B - 1) \cdot N_{\text{while}} \cdot [\text{time}_{\min \mathcal{H}} + N_{\text{SP}} \cdot ((N_{\text{RK}} \cdot N_y \cdot N_{g_t} \cdot C_{\text{RK}}) + C_{\varpi > 3})]$ | (5.66) |
| | Evaluate: | $N_{\text{TPBVP}} \cdot N_p \cdot (N_B + C_{\text{Eva}})$ | (5.67) |
| Broyden method: | | $(N_{\text{TPBVP}} - 1) \cdot (N_x^2 + N_x) \cdot C_{\text{Broyden}}$ | (5.68) |

The values of $\text{time}_{\min l}$ and $\text{time}_{\min \mathcal{H}}$ depend on what kind of problems they are and how they are solved. Without Lemma 1, both problems are treated as NLP problems, due to the nonlinear equality constraints. A broadly available solution method in commercial software, e.g. Matlab, is SQP. A SQP method was examined in [147], which showed that it takes no worse than $O(\epsilon^{-2})$ iterations to solve a constrained nonlinear problem to ϵ accuracy, where $\epsilon \in (0,1)$ is the accuracy of the first-order necessary conditions for optimality [148]. Given the complexity of quadratic programming at each iteration to solve a quadratic subproblem is $O(n^3)$, the complexity of the whole solution is approximately $O(n^3 \epsilon^{-2})$, where n is the number of the inputs. Please be aware, this is not a strict evaluation, since the approximation of the constraints has not been discussed. With Lemma 1, both problems are treated as convex optimization problems in a standard form, which can be solved by the interior-point method, an efficient method to solve convex optimization problems (see [63, pp. 561-620] for more information of the method). It was reported in [149] that the complexity of using it to solve a convex optimization problem to ϵ accuracy is $O(n^{3.5} \log(1/\epsilon))$. The default value of ϵ in Matlab is 1×10^{-6} [150]. Given the small number of inputs of the optimization problem analysed in this Chapter, Lemma 1 reduces the complexity. Lemma 2 and Lemma 3 modify the inequality constraints so that the feasible set of the inequality constraints becomes smaller, through which the interior-points method requires less iterations⁵⁰ [63, pp. 592-594]. The

⁵⁰ Assume that the feasible set of the inequalities is contained in a Euclidean ball of radius R . The complexity is

reduction of $\text{time}_{\min l}$ and $\text{time}_{\min \mathcal{H}}$ through Lemma 2 and Lemma 3 is less significant than Lemma 1.

The first two most dominant parts in Table 5.2 are those involving the minimization and the RK integration scheme. In big-O notation, their time complexities⁵¹ are $O(N_{\text{TPBVP}} \cdot N_p \cdot N_B \cdot N_{\text{while}} \cdot n^3 \cdot \epsilon^{-2})$ and $O(N_{\text{TPBVP}} \cdot N_p \cdot N_B \cdot N_{\text{while}} \cdot N_{\text{RK}} \cdot N_y \cdot N_{g_c})$, respectively. With Lemma 1, the former is reduced to $O(N_{\text{TPBVP}} \cdot N_p \cdot N_B \cdot N_{\text{while}} \cdot n^{3.5} \cdot \log(1/\epsilon))$.

N_{TPBVP} depends on the algorithm applied to solve the TPBVP and the initialization strategy. They are the Broyden method and the Bang-Bang controls inspired strategy. A theoretical analysis is difficult to perform. Their effects will be evaluated in a numerical simulation in Section 6.3.1.

5.5 Summary

This chapter has adopted the HMP and has developed a predictive EMS embedded with the HMP solution method. The developed predictive EMS is schematically illustrated in Figure 5.10, in which the major components for the algorithm are presented. The Broyden method has been applied for the first time for the solution of the HMP in the context of EMSs. The shooting process is innovatively initialized by a strategy inspired by Bang-Bang controls. Importantly, the instantaneous Hamiltonian has been mathematically analysed, which produces Lemma 1, 2 and 3. The behaviour of the predictive EMS is to be evaluated in Chapter 6.

It should be pointed out that the algorithms in the predictive EMS are suitable to multidimensional problems. On the other side, the works that used IMs that concern 1D problems—which are in most cases in the literature, as mentioned in Section 2.4.2 and Section 3.2.1.2—are difficult, if not impossible, to be extended to higher dimensional problems.

The HMP solution method involves the minimization of the instantaneous Hamiltonian, a non-linear function, at each time step. The foremost advantage of showing the minimization is a convex optimization problem is that the problem can be solved reliably and efficiently with the interior-point method or other special methods for convex optimization [63, preface]. Furthermore, most of these efficient and available algorithms require a standard formulation. Unfortunately, the problem (5.33) turns out not to be a convex optimization problem, let alone one in a standard form. Lemma 1 resolves the difficult situation with the prerequisite: the power losses of both EMs and both STs are convex in the direction of torque. The power losses in this thesis are the simulation data provided by the project partners. They fulfil the prerequisite (see Proof of Lemma 1).

dependent on $\log R$.

⁵¹ Big-O notations ignore constants and lower order terms.

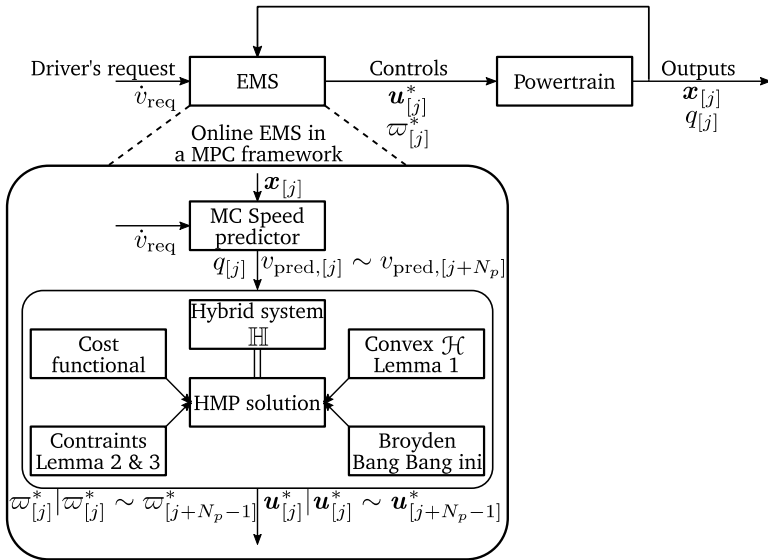


Figure 5.10: Structure of the predictive EMS embedded with the HMP solution method, supplemented by Lemma1, 2 and 3. TPBVP solved by the Broyden method with the Bang-Bang controls inspired initialization.

The convexity of the power losses might not hold if, among other possibilities, measured power losses are used, due to measure errors etc. It can happen when either angular velocity or torque is close to zero. One possible way to use the Lemmas would be separating the optimization domain.

The time complexity of the algorithm of the predictive EMS is analysed. It is noticed that Lemma 1, 2 and 3 have reduced the worst-case complexity of the algorithm. They can be transformed to the proofs and the treatments for the optimization problems with similar settings in other applications and studies, even those that are not relevant to EMSs.

6 Simulation Results of the Predictive Energy Management Strategy

A predictive EMS embedded with the HMP solution method is developed in Chapter 5. The time step h is set to 1 s and the time step during shift processes h_{sp} is set to 0.01 s. This chapter discusses the simulation results of the developed EMS.

The first part introduces two reference EMSs: (a) EMS naive optimal, an EMS that minimizes instantaneous energy consumption but ignores shift processes and (b) EMS global optimal, a global optimal provided by the DP, given the speed profile *a priori*. The EMS global optimal determines the theoretical optimum for the HOCP, for which the predictive EMS strives. Note that both reference EMSs are identical, if the shift processes are neglected in evaluation, which is used in [13].

The second part investigates the solution process of the predictive EMS. In particular, it sheds light on the process of solving HOCP through minimizing the instantaneous Hamiltonian. To showcase the advantages of the HMP in the context of EMS, its solutions at several representative driving conditions are compared to those of EMS naive optimal. Furthermore, the iteration process is presented. The convergence behaviour of `solveHMP` with the Broyden method is compared to a commonly available Quasi-Newton method that requires little effort to implement.

The third part shows the result of driving cycle simulation. The WLTC is chosen as the speed profile for the simulation, which is widely used as a standard driving cycle for evaluating EMSs. The predictive EMS is compared to the reference EMSs. Additionally, the predictive EMS is studied w.r.t. accuracy of predicted speed and length of the prediction horizon.

6.1 Reference Energy Management Strategies

Both reference EMSs are presented in the following subsections.

6.1.1 EMS Naive Optimal

A reference EMS minimizes the instantaneous cost function (5.7) subject to the constraints on the EM torques and the angular velocities (5.13)-(5.17), the equality constraint on the total output torque (4.16) and the dynamics of the continuous states (4.14)-(4.18). It assumes that a shift process takes place instantly without energy consumption or recuperation, which limits the discrete state to $q \in \{q_1, q_2, q_3\}$. With this setting, the EMS naive optimal has two DoFs and two inputs: the vehicle speed and the total output torque. Optimization is performed over the whole input space to obtain the look-up tables of the target gear positions and the EM torques. Figure 6.1 shows the look-up table of the target gear positions.

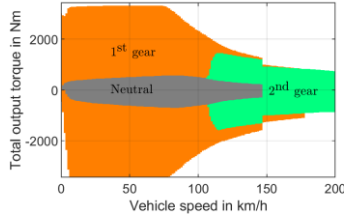


Figure 6.1: Target gear position look-up table

6.1.2 EMS Global Optimal

For a given driving cycle, the forward discrete DP⁵² is implemented to solve the HOCP (4.51) under the constraints (4.56)-(4.63). The length of time step is set to be 1 s. At each time step $t_{[k]}$, the discrete control $\varpi_{[k]}$ is enumerated in the admissible discrete control space $\mathcal{B}_{q_{[k]}}$ (see Table 4.2). For the discrete control representing the fixed gear positions, $u_{1,[k]}$ represents the EM1 torque at $t_{[k]}$. For the discrete control representing the shift processes, $u_{1,[k]}$ represents the EM1 torque after the shift process in the new gear position, i.e. $T_{1,\text{new}}$ in (4.42). The torques during the shift processes are predefined, as shown in Section 4.4.2. The admissible continuous space of $u_{1,[k]}$, i.e. $[T_{1,\min}(\omega_{1,[k]}), T_{1,\max}(\omega_{1,[k]})]$, is discretized to generate a grid \mathcal{G}_{u_1} of 1000 points, so that the increment on the grid is smaller than 0.1 Nm. $u_{1,[k]}$ is enumerated on the discretization grid $\mathcal{G}_{u_{1,[k]}}$. $u_{2,[k]}$ can be determined according to $u_{1,[k]}$, the gear position G and the constraint (4.16).

6.2 Solution Process of the Predictive EMS

Given a combination of current speed and acceleration, the predicted speed profile in the prediction horizon, $N_p = 10$ in this subsection, are generated by **predMC** starting from $t = 0$ s, as described in Section 5.1. Several examples of representative driving conditions are presented in Figure 6.2, in which (a) and (b) are the vehicle operating points from WLTC to represent launch with small acceleration and deceleration at low speed, respectively. A speed profile of constant high acceleration is formed in Figure 6.2(c). They are chosen to demonstrate the properties of the HMP

⁵² To solve an OCP with the discrete DP is to solve the Bellman equation (2.48)-(2.50). Due to its recursive nature, it is natural to perform the calculation in backwards time direction, i.e. in the order of $t_{[N_t]}, t_{[N_t-1]} \dots t_{[0]}$. As a result, states are calculated with backwards integration. In the case of free terminal conditions but fixed initial conditions, e.g. the gear position is unknown at $t_{[N_t]}$ but known at $t_{[0]}$, it is tricky to assign reasonable values for states at $t_{[N_t]}$ to obtain the global optimal. To avoid such inconvenience, the HOCP in question is solved in forward time direction, which is equivalent to using backward discrete DP to solve the HOCP with reversed speed profile and switched boundary conditions. Since its equivalency, the Bellman equation for the forward DP is omitted.

solution process that reaches different solutions compared to the EMS naive optimal.

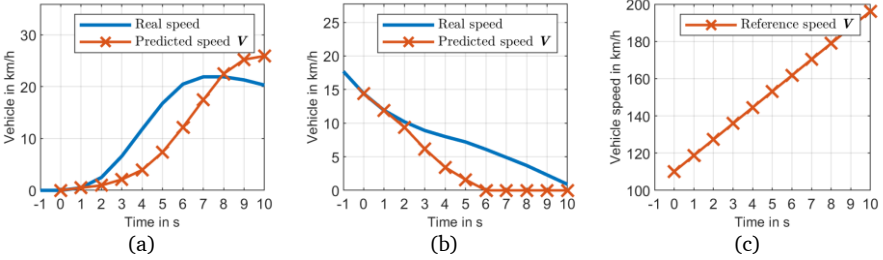


Figure 6.2: Predicted speed of individual solutions. (a) $t = 512$ s in the WLTC. (b) $t = 376$ s in the WLTC. (c) high speed and moderate acceleration.

Given the predicted speed $\mathbf{V} := [v_{\text{pred},[0]}, v_{\text{pred},[1]}, \dots, v_{\text{pred},[N_p]}]$, **solveHMP**, shown in Figure 5.4, solves the TPBVP transformed from the HOCP iteratively. At an iteration, **shootHMP**, shown in Figure 5.3, determines the discrete control sequence and the continuous control functions by minimizing the instantaneous Hamiltonian through the prediction horizon. The following subsections discuss this process.

6.2.1 Minimizing the Instantaneous Hamiltonian

This subsection discusses the difference between minimizing the instantaneous Hamiltonian and minimizing the instantaneous cost with the examples from Figure 6.2, and does minimizing the instantaneous Hamiltonian minimize the cost functional in the whole time interval.

Launch with small acceleration

Figure 6.3 presents such a process at the 1st second of an intermediate iteration of **solveHMP** solving the HOCP given the predicted speed profile in Figure 6.2(a). Starting with the neutral gear position, i.e. $q(t_{0,p}) = q_3$, the powertrain can stay in the neutral or shift to the 1st or the 2nd gear position, i.e. the admissible discrete control set $\mathcal{B}_{q_3} = \{q_3, q_8, q_9\}$. Figure 6.3(a) shows the angular velocities of both EMs for each admissible discrete control. Since the ST1 is single-speed, the discrete control does not change ω_1 . ω_2 evolves differently according to the different discrete controls. Due to low vehicle speed, the shift durations in q_8 and q_9 at this instance are short. Thereafter, $\mathbf{u}_q^* = (T_{1,q}^*, T_{2,q}^*)^T$ are determined through minimizing \mathcal{H}_q and are presented in Figure 6.3(b). As shown in Section 5.2.5, it is equivalent to minimizing the instantaneous cost with a fixed q . (5.19) determines the costates, which are illustrated in Figure 6.3(c). Though ω_1 evolves the same regardless of the discrete control, λ_1 evolves differently, due to the different torques. The cost function and the Hamiltonian as well as their cumulated values are presented in Figure 6.3(d) and (e). The shift processes of q_8 and q_9 consume energy and, therefore, l_{q_8} and l_{q_9} , as well as their

cumulated values are higher than those of q_3 in the first 0.2 second. However, the system consumes less energy in the 1st and the 2nd gear position for the rest of the time, so that the cumulated cost of q_3 at the end of 1st second is the highest among the three. An EMS that minimizes the instantaneous cost, i.e. the electric power, may it consider shift processes or not, chooses q_B . Interestingly, under the HMP, the predictive EMS chooses q_3 , since the resulting Hamiltonian of q_3 , i.e. $\mathcal{H}_{q_3} = l + \lambda_{q_3}^T \cdot f_{q_3}$ (see (5.18) in Section 5.2.2), is averagely and cumulatively the lowest.

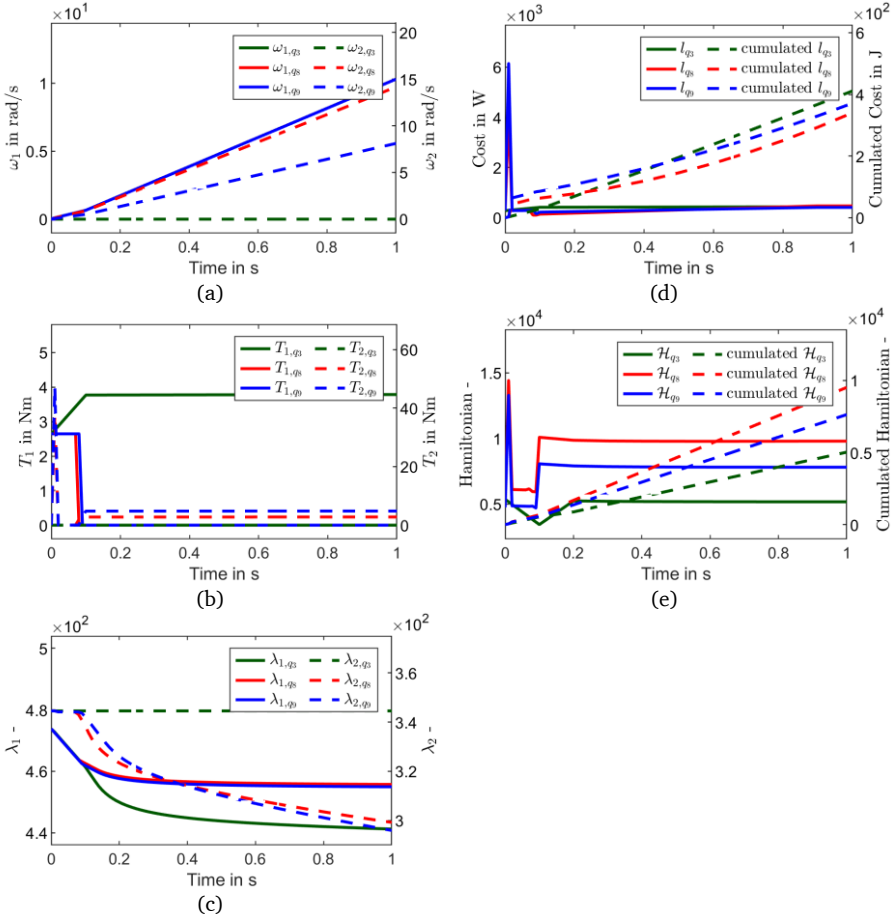


Figure 6.3: The instantaneous Hamiltonian in the 1st second: (a) angular velocities, (b) torques, (c) costates, (d) cost function and (e) Hamiltonian.

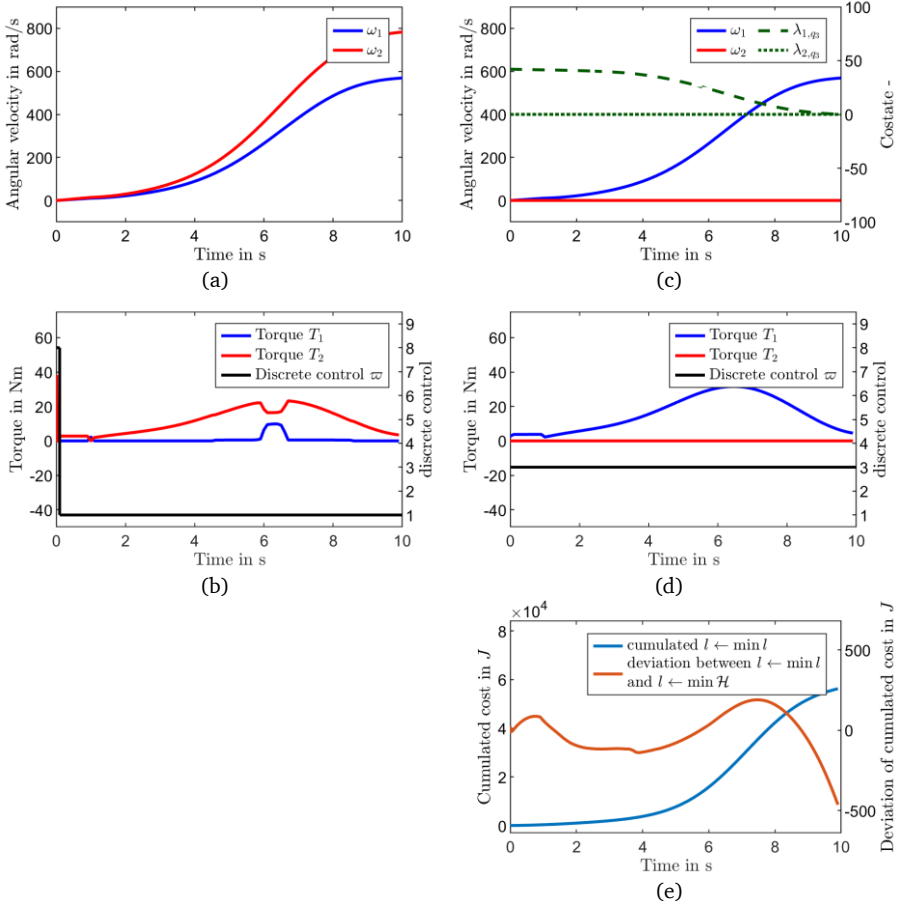


Figure 6.4: Comparison of the solution by minimizing the instantaneous cost, (a) and (b), to the one by minimizing the instantaneous Hamiltonian, (c) and (d). (a) and (c): states and costates. (b) and (d): continuous and discrete controls. (e) Deviation of cumulated cost.

One might ask: is minimizing the instantaneous Hamiltonian beneficial? Figure 6.4 compares the solution of minimizing the instantaneous cost to the one generated by the HMP over the whole prediction horizon. Figure 6.4(a) shows the continuous states of the former solution. To achieve the lowest cost at each time step the system remains in q_1 , after it shifts from the neutral to the 1st gear position, i.e. q_8 at the beginning of the time interval. The corresponding controls are presented in Figure 6.4(b). On the other hand, under the HMP, the system remains in q_3 over the time

interval, as the discrete control sequence in Figure 6.4(d) shows. The state and the costate trajectories are presented in Figure 6.4(d). **solveHMP** finds a solution, as the costates in Figure 6.4(c) evolve to zero and fulfil the transversality condition. Figure 6.4(e) compares the cumulated cost over the time interval. The blue curve represents the value of minimizing the instantaneous cost and the brown curve shows the deviation between both solutions. A negative deviation at $t = 10$ s means that the solution by the HMP reduces cumulated cost over the whole time interval. The HMP manages to minimize the integral, though the cost during the time interval are not always the minimal, which is indicated by a positive deviation in Figure 6.4(e).

It is worth mentioning that λ plays an important role during the process. Only when $\lambda^*(\cdot)$ follows its dynamics in (2.53) and meets the transversality condition in (2.55), the HMP locates a minimal for the HOCB⁵³. Figure 6.5 shows the Hamiltonian in the 1st second with negative $\lambda_1(t = 0)$ and $\lambda_2(t = 0)$ values, which are randomly initialized instead of using the initialization inspired by Bang-Bang control developed in 5.2.4. $\mathbf{x}_q(\cdot)$, $\mathbf{u}_q(\cdot)$ and $l_q(\cdot)$ are the same as Figure 6.3(a), (b) and (d). At this iteration, **shootHMP** chooses q_B for the 1st second, since \mathcal{H}_{q_B} is the smallest, which leads to the same controls as Figure 6.4(b). Section 6.2.2 will present the results of the iteration process of **solveHMP**, through which $\lambda^*(\cdot)$ is found.

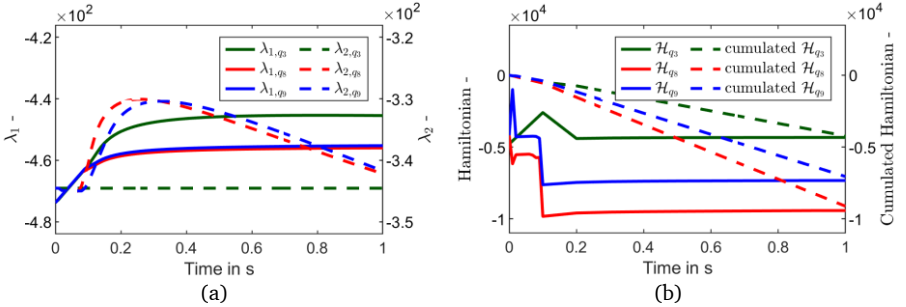


Figure 6.5: Hamiltonian given different costates in 1st second. (a) costates. (b) Hamiltonian.

Deceleration at low speed

For the predicted speed profile in Figure 6.2(c), the solutions generated by the EMS naive optimal and by the predicted EMS are presented in Figure 6.6. The system starts with the 1st gear position. Figure 6.6(b) shows that the EMS naive optimal requested 2 shifts, the 1st to the neutral gear position (q_6) and the neutral to the 1st gear position (q_8). On the other hand, the predicted EMS requests a constant gear position during the time interval, as Figure 6.6(d) shows. It has the vehicle consume 326 J less energy (3.5 %).

⁵³ See [16, pp. 184-189] and [16, pp. 228-236], if the reader wishes to understand the mathematical role of costates and Hamiltonian.

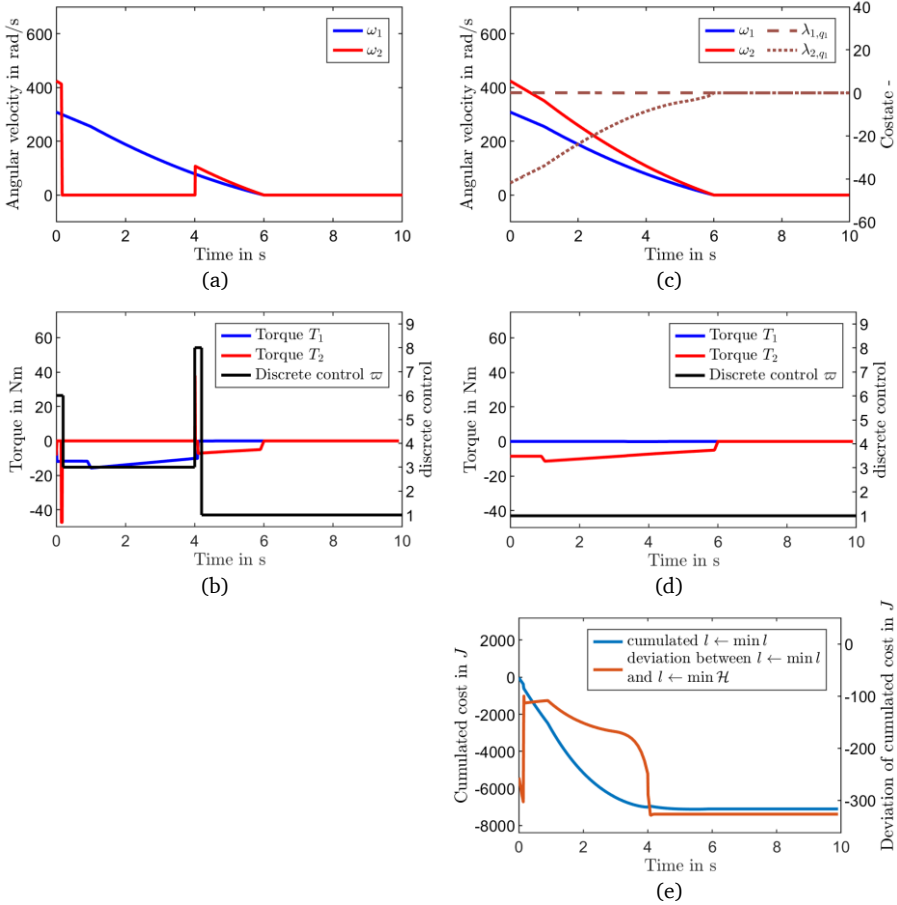


Figure 6.6: Comparison of the solution by minimizing instantaneous cost, (a) and (b), to the one by minimizing instantaneous Hamiltonian, (c) and (d). (a) and (c): states and costates. (b) and (d): continuous and discrete controls. (e) Deviation of cumulated cost.

Moderate acceleration at high speed

The requested acceleration in the above mentioned two predicted speed profiles is within the range of the admissible output torque of the powertrain. The example shown in Figure 6.2(c) includes a combination of speed and acceleration outside that range.

Starting with the 1st gear position, under the EMS naive optimal, the system shifts to the 2nd gear position (q_4) in the first second, as shown in Figure 6.7(a) and (b), for the combination

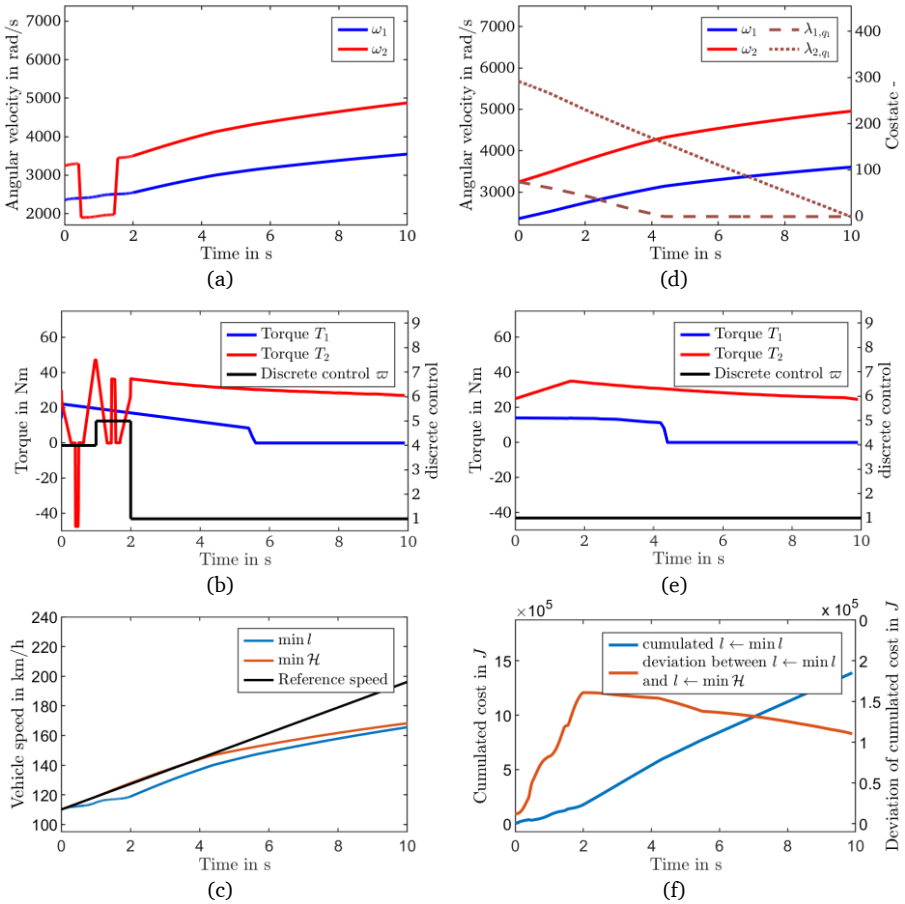


Figure 6.7: Compare the solution by minimizing instantaneous cost, (a) and (b), to the one by minimizing instantaneous Hamiltonian, (d) and (e). (a) and (d): states and costates. (b) and (e): continuous and discrete controls. (c) deviation of vehicle speed. (f) deviation of cumulated cost.

of speed and output torque enters the “2nd gear position” area of the look-up table in Figure 6.1. During the shift process, T_1 is limited by its upper bound that lowers as ω_1 increases (see Figure B.1), and cannot compensate the change of T_2 to accelerate the vehicle as requested. Consequently, controlled by the EMS naive optimal, the vehicle speed (blue curve in Figure 6.7(c)) lags after the shift process in the first second. In order to restore the vehicle speed, the acceleration is modified to a higher level, and the combination of speed and output torque moves upwards in the look-up

table in Figure 6.1. The 1st gear position is selected as the target gear. Same to the first shift process, the second shift process causes a speed lag. The EM1 continuous to provide torque until ω_1 reaches its upper bound (5.15). On the other hand, under the predictive EMS, the system stays in q_1 , as shown in Figure 6.7(d) and Figure 6.7(e). The HMP generates a solution with a higher vehicle speed that is closer to the reference speed. It consumes more energy, as shown in Figure 6.7(f).

The EMS naive optimal does not consider the information of the system dynamics, which can cause consecutive shifts between two gear positions and worsen the drivability of the vehicle.

6.2.2 Convergence of the Solution

The algorithm `solveHMP` initializes the solution, evaluates the outputs of `shootHMP` iteratively through the Broyden method and determines the optimal control.

For the solution process of the speed profile in Figure 6.2(a), Figure 6.8(a) shows the iterations of the error of the transversality condition on a logarithmic scale, in which the subscript denotes the index of elements in a vector. The absolute values $|\Phi_1|$ and $|\Phi_2|$ shrink to 2.44×10^{-14} and 0.63×10^{-14} with 7 iterations, respectively. The convergence behaviour for this solution is comparable to a fast 1-D method called “Pegasus” reported in [71] that finds the root of $x^3 + 1 = 0$ with an error of 10^{-14} magnitude with 9 iterations. Figure 6.8(b) shows the indexed iterations with the initial costates on the x- and y-axis. The z-axis shows the Euclidean norm of the error $\|\Phi\|$ on a logarithmic scale. The first two guesses, marked green, are generated by `genInitialCostate` (see Section 5.2.4), and the rest is updated by the Broyden method in `solveHMP`. The norm of the error at the final iteration, marked red, reaches a magnitude of 10^{-14} .

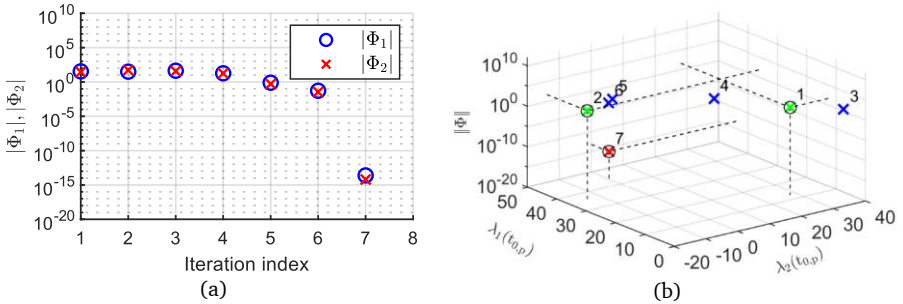


Figure 6.8: Convergence behaviour for the speed profile in Figure 6.2(a). (a) The error of the transversality condition at each iteration. (b) The initial costates vs. the Euclidean norm of errors at each iteration

The error does not necessarily have to converge to a magnitude of 10^{-14} . As described in `solveHMP` in Figure 5.4, the limit for convergence is positively correlated to the norm of $\hat{\lambda}_{[0],1}$ and $\hat{\lambda}_{[0],2}$, i.e. the first two proposed costates. Figure 6.9 shows the iteration of errors of the solution

process for the speed profile in Figure 6.2(b). The error reaches a magnitude of 10^{-15} at the 10th iteration. Since the 6th iteration, the discrete and continuous control functions presented in Figure 6.6(d) is found. At the 8th iteration, the error meets the convergence limit, i.e. $\|\Phi\| < \varepsilon$. Figure 6.10 shows the iteration of errors of the solution process for the speed profile in Figure 6.2(c). $\|\Phi\|$ reaches a magnitude of 10^{-11} at the 6th iteration.

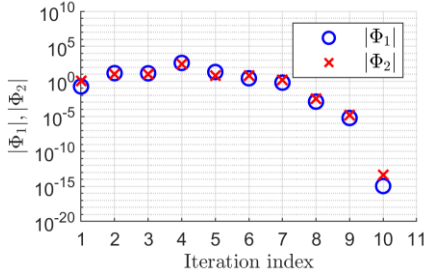


Figure 6.9: $|\Phi_1|$ and $|\Phi_2|$ of the solution process for the speed profile in Figure 6.2(b).

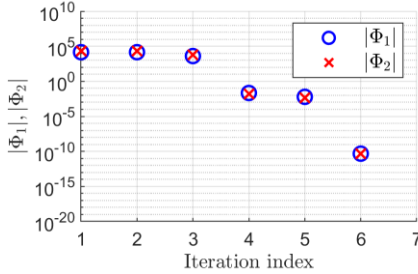


Figure 6.10: $|\Phi_1|$ and $|\Phi_2|$ of the solution process for the speed profile in Figure 6.2(c).

6.2.3 Convergence Behaviour Compared to SQP

As mentioned in Section 2.4.1, Quasi-Newton methods are suitable for the TPBVP. SQP, one of Quasi-Newton methods, is available in various software, e.g. Matlab and Python. It is implemented in an EMS in Matlab with its standard optimization toolbox⁵⁴. Figure 6.11 presents the structure of the EMS, in which the only difference is that $\hat{\lambda}_{[0]}$ in `solveHMP` (see Figure 5.4) is iteratively proposed by SQP instead of the Broyden method. The convergence behaviour of `solveHMP` with

⁵⁴ Solving the TPBVP is considered minimizing the error of the transversality condition (5.27), $\min_{\lambda_{[0]}} \|\Phi\|$, with the SQP active-set algorithm proposed in [151, 152]. The minimization is initialized with point $(0,0)^T$ and executed with default setting in Matlab.

SQP is compared to the one developed in Chapter 5, i.e. **solveHMP** with the Broyden method.

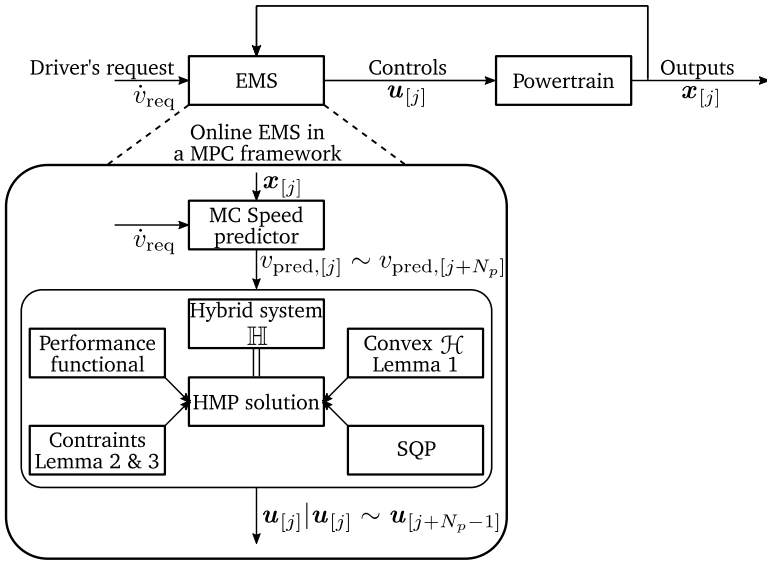


Figure 6.11: Structure of the predictive EMS embedded with HMP solution method, whose TPBVP solved by SQP.

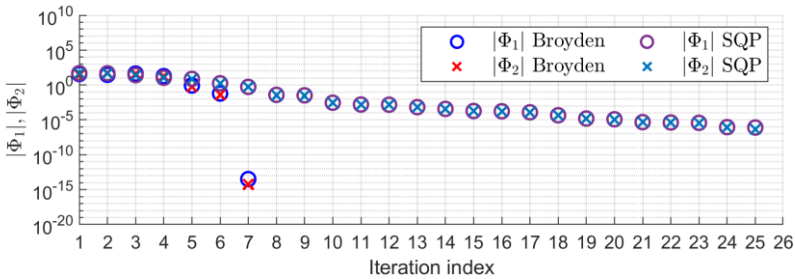


Figure 6.12: Comparison of the iterations of error

For the same predicted speed profile as in Figure 6.2(a), convergence of the solution can be reached through **solveHMP** with SQP. Figure 6.12 compares the iterations of error of both algorithms. The error of the transversality condition of the solutions proposed by **solveHMP** with SQP shrinks gradually and steadily. It meets the convergence limit, i.e. $\|\Phi\| < \epsilon$, in 25 iterations, which is slower and less accurate than **solveHMP** with the Broyden method.

6.3 Driving Cycle Simulation

Driving cycle simulations with the WLTC (see Figure 5.2) are performed. The vehicle model follows the dynamics of the hybrid system described in Section 4.4. Its EMS is called once every 1 s. The length of the prediction horizon is 10 s and that of control horizon is 1 s.

6.3.1 Iteration Statistics

At each time point in the WLTC, **predictiveEMS** developed in Section 5 is carried out. Figure 6.13 summarizes the iteration number, i.e. N_{TPBVP} in Section 5.4, that each solution from **solveHMP** takes to converge. The solutions at standstill are discarded. With the Broyden method and the initialization developed in Section 5, 1429 out of 1575 solutions (90.6 %) require less than or equal to 7 iterations. The convergence behaviour in the WLTC remains unchanged with different value of coefficient β in (5.8), since the speed calculated in the **solveHMP** is close to the constrained predicted speed. $\beta = 1$ is chosen for the simulation study in the rest of this work. On the other hand, with SQP, the predictive EMS in Figure 6.11 requires at least 20 iterations to converge.

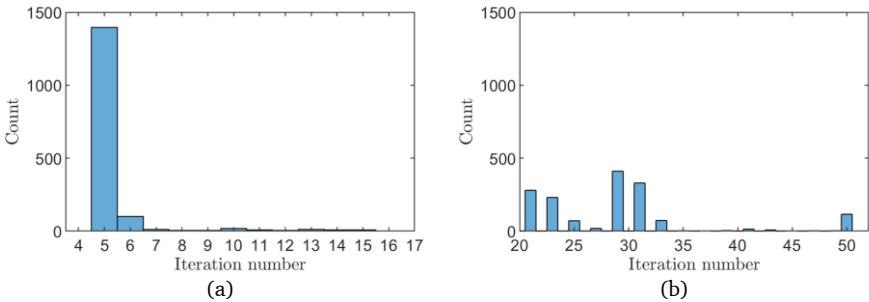


Figure 6.13: Iteration numbers for solutions to converge. (a): solveHMP with the Broyden method. (b): solveHMP with the SQP.

6.3.2 Comparison to EMS Naive Optimal

This subsection compares the predictive EMS to the EMS naive optimal. The comparison focuses on the energy consumption and the shift count in the driving cycle.

Overall evaluation

Figure 6.14 compares the discrete state sequences of the EMS naive optimal, upper half, and the predictive EMS, lower half. Each discrete state is marked with a colour listed in the legends. In the WLTC, the reference EMS requests in total 123 shifts, while the predictive MPC requests in total 45 shifts. The energy consumption of the EMS naive optimal in the WLTC driving cycle simulation is 11.55 kWh/100 km, while the value of the predictive EMS is 11.53 kWh/100 km. The predictive

EMS reduces 0.26 % energy consumption and 63.41 % shift count.

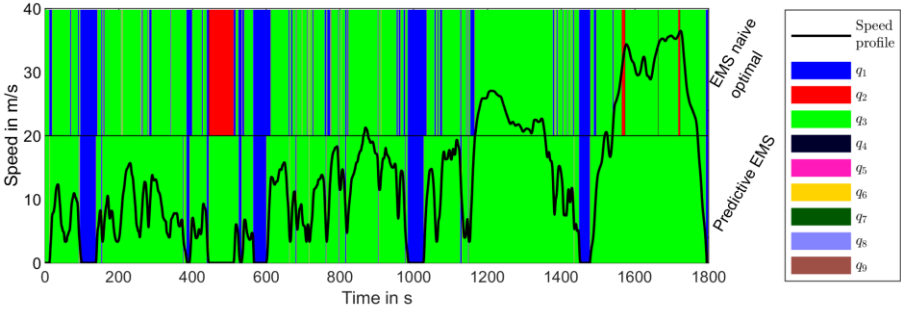


Figure 6.14: Discrete state sequences of the EMS naive optimal and the predictive EMS

Although the predictive EMS has reduced a large number of shifts successfully compared to the EMS naive optimal, 10 unnecessary shifting instances happen, at which the ST2 shifts between two gear positions within 2 seconds. They are caused by the change of the sign of requested acceleration, which is not reflected by the predicted speed profile. Figure 6.15 shows two examples.

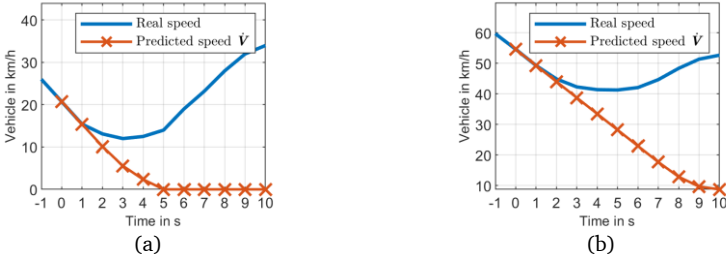


Figure 6.15: Predicted speed vs. real speed at (a) 663 s and (b) 797 s.

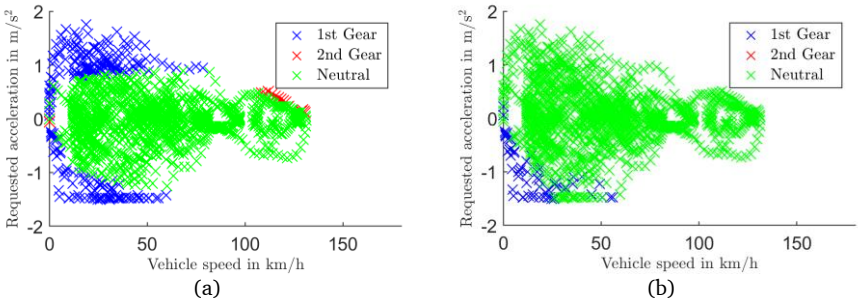


Figure 6.16: Gear distribution of (a) the EMS naive optimal and (b) the predictive EMS.

Gear distribution and EM operating points

Figure 6.16 compares the gear distribution of both EMSs in the WLTC. Based on Figure 6.14 and Figure 6.16, it is easy to notice that the ST2 primarily stays in the neutral gear position, when it is controlled by the predictive EMS. It only shifts to the 1st gear position under negative acceleration.

The vehicle, under the control of the EMS naive optimal, shifts to the 2nd gear position at high speed and positive acceleration. Figure 6.17 shows the control and the resulting energy consumption of both EMSs between 1560 s and 1580 s of WLTC, during which the ST2 appears in the 2nd gear position. Energy is consumed and is recuperated, when the ST2 shifts to the 2nd gear position and back to the neutral gear position in Figure 6.17(c), respectively. It consumes more energy compared to the predictive EMS, which is shown by a negative deviation in Figure 6.17(b).

Figure 6.16(a) shows that, under the EMS naive optimal, the ST2 shifts to the 1st gear position, when either the absolute value of acceleration is higher than or equal to 1 m/s^2 or the speed is close to zero. The predictive EMS, on the other hand, requests the same, when the acceleration is lower than or equal to -1 m/s^2 or the speed is close to zero. Figure 6.18 summarizes the operating

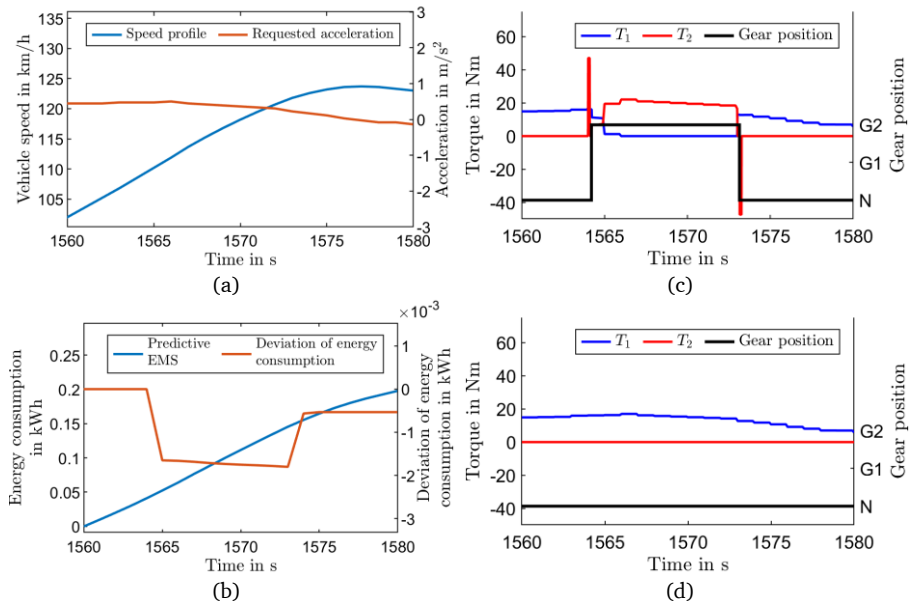


Figure 6.17: Control and energy consumption between 1560 s and 1580 s. (a) Speed and acceleration. (b) Energy consumption and its deviation. (c) Torques and gear positions of the EMS naive optimal. (d) Torque and gear positions of the predictive EMS.

duration of both EMs at different operating points. On the X and Y axes, the power losses of EMs are illustrated. It can be noticed that, with the EMS naive optimal, the EM2 operates at the operating points with 20 Nm and -20 Nm longer than the case with the predictive EMS, which correspond to the areas of 1st gear position above 1 m/s² and below -1 m/s² in Figure 6.16(a).

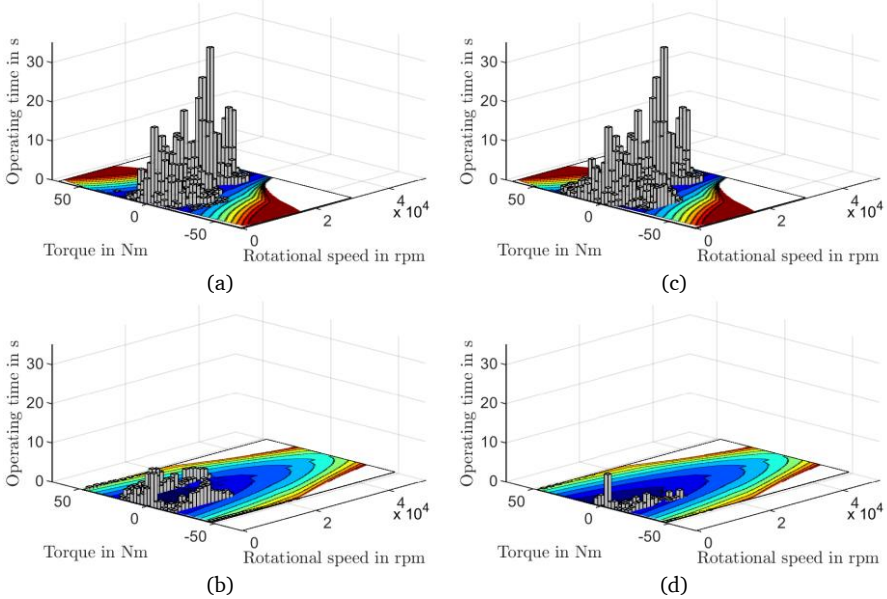


Figure 6.18: Operating duration of both EMs. (a) and (b): the EM1 and the EM2 of the EMS naive optimal. (c) and (d): the EM1 and the EM2 of the predictive EMS.

Figure 6.19 shows the control and the resulting energy consumption of both EMSs between 255 s and 295 s of WLTC, during which the ST2 appears in the 1st gear position. With the EMS naive optimal, The ST2 shifts back and forth between the neutral and the 1st gear positions, as shown in Figure 6.19(c). The energy saved in the 1st gear position and the recuperated energy does not compensate the energy consumed for shifting.

6.3.3 Comparison to the EMS Global Optimal

This subsection compares the predictive EMS to the EMS naive optimal.

Overall evaluation

Figure 6.20 compares the discrete state sequences of the EMS global optimal, upper half, and the predictive EMS, lower half. The sequences of both EMSs largely coincide with each other. Thanks to the information of the speed profile as *a priori* and the hybrid formulation including the shift

processes, the EMS global optimal only requests in total 27 shifts in WLTC, which is lower than that of the predictive EMS. The total energy consumption of the EMS global optimal in the WLTC driving cycle simulation is 11.51 kWh/100 km, while the value of the predictive EMS is 11.52 kWh/100 km. The predictive EMS falls short by 0.018 kWh/100 km (0.18 %) compared to the global optimal.

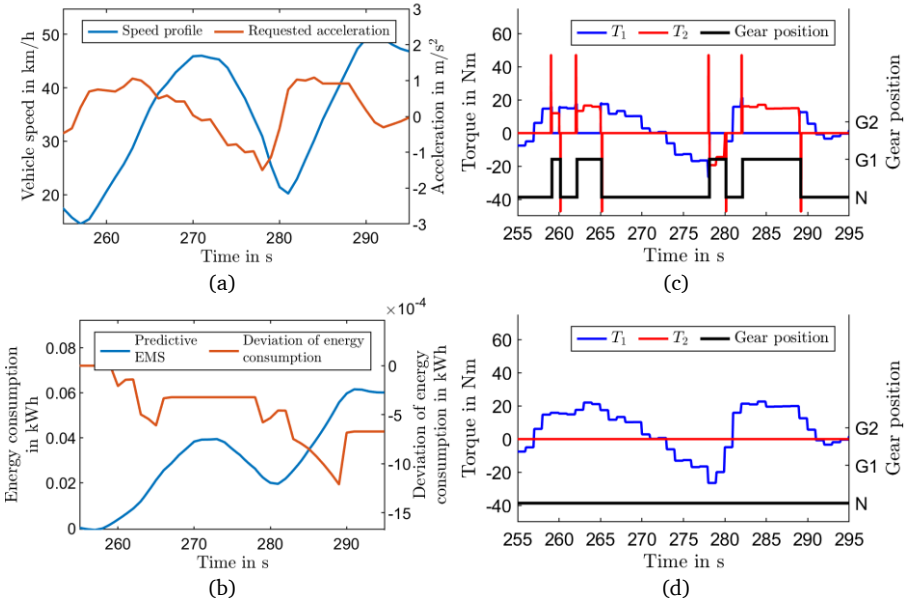


Figure 6.19: Control and energy consumption between 255 s and 295 s. (a) Speed and acceleration. (b) Energy consumption and its deviation. (c) Torques and gear positions of the EMS naive optimal. (d) Torque and gear positions of the predictive EMS.

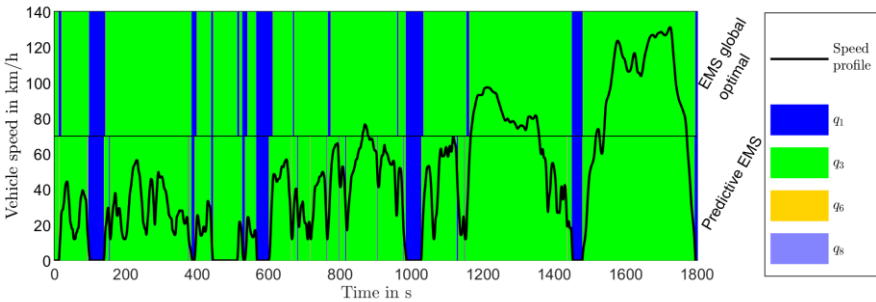


Figure 6.20: Discrete state sequences of the EMS global optimal and the predictive EMS.

Gear distribution

After comparing the gear distribution of the EMS global optimal in Figure 6.21 and that of the predictive EMS in Figure 6.16(b), it can be noticed that the EMS global optimal requests the 1st gear position with acceleration larger than 1 m/s^2 , e.g. near 765 s, 960 s and 1160 s in Figure 6.20.

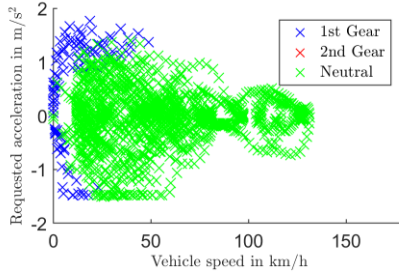


Figure 6.21: Gear distribution of the EMS global optimal.

Figure 6.22 shows the control and the resulting energy consumption of both EMSs between 1150 s and 1170 s of the WLTC. With the EMS global optimal, the ST2 appears in the 1st gear position for ca. 8 s, as shown in Figure 6.22(c). The deviation in Figure 6.22(b) between 1153 s and 1161 s declines, which indicates that the vehicle consumes more energy in the 1st gear position (controlled by the EMS global optimal) than in the neutral gear position (controlled by the predictive EMS). However, with the control shown in Figure 6.22(c), the EM2 recuperates much more energy at the shift from the 1st to the neutral gear position than it consumes at the shift from the neutral to the 1st gear position, since the former shift process has a larger angular velocity difference. The energy consumed and recuperated during both shift processes are compared against each other in Figure 6.23. As a result, the predictive EMS consumes more energy between 1150 s and 1170 s. The predictive EMS falls short in this case, due to its less accurate speed prediction. At 1153 s when the EMS global optimal requests an upshift based on the WLTC as *a priori*, a predicted speed profile below the real speed profile in the prediction horizon is generated for the predictive EMS, which is shown in Figure 6.24. The same situation happens at 765 s and 960 s.

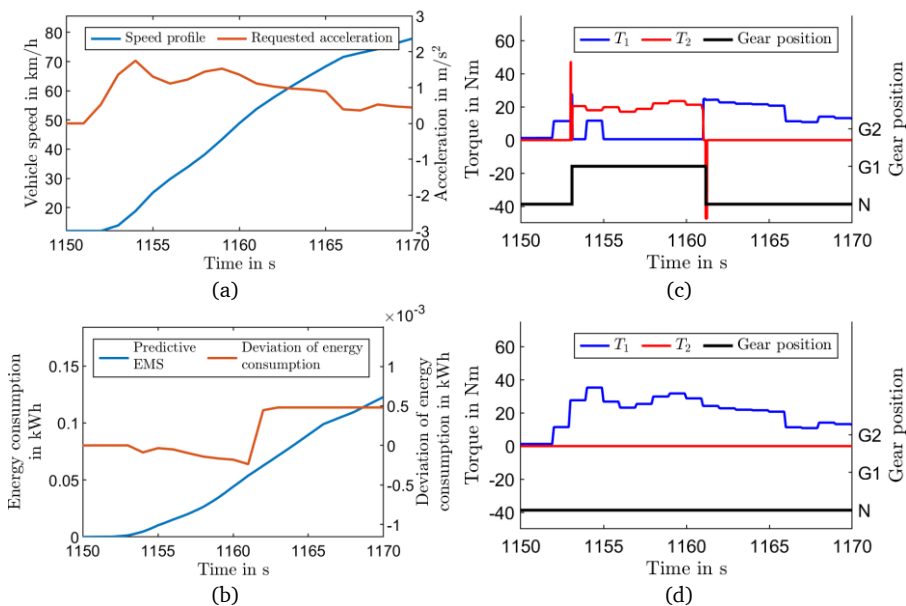


Figure 6.22: Control and energy consumption between 1150 s and 1170 s. (a) Speed and acceleration. (b) Energy consumption and its deviation. (c) Torques and gear positions of the EMS naive optimal. (d) Torque and gear positions of the predictive EMS.

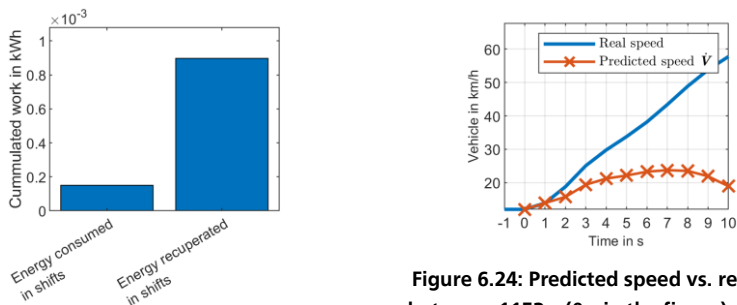


Figure 6.23: Energy consumed and recuperated during shifts between 1150 s and 1170 s.

Figure 6.24: Predicted speed vs. real speed between 1153 s (0 s in the figure) and 1163 s (10 s in the figure).

6.3.4 Influence of the Prediction Accuracy

Sections 6.3.2 and 6.3.3 have shown the negative influence of inaccurate prediction on the control with examples. To investigate the influence of the accuracy of the predicted speed profile, this

subsection replaces the MC predictor in the predictive EMS in Figure 5.10 with a zero-error prediction model, as shown in Figure 6.25. At $t_{[j]}$, the speed from the driving cycle are considered the predicted speed in the prediction horizon, as

$$\left[v_{\text{pred},[j]}, v_{\text{pred},[j+1]}, \dots, v_{\text{pred},[j+N_p]} \right] = \left[v_{\text{cycle},[j]}, v_{\text{cycle},[j+1]}, \dots, v_{\text{cycle},[j+N_p]} \right]. \quad (6.1)$$

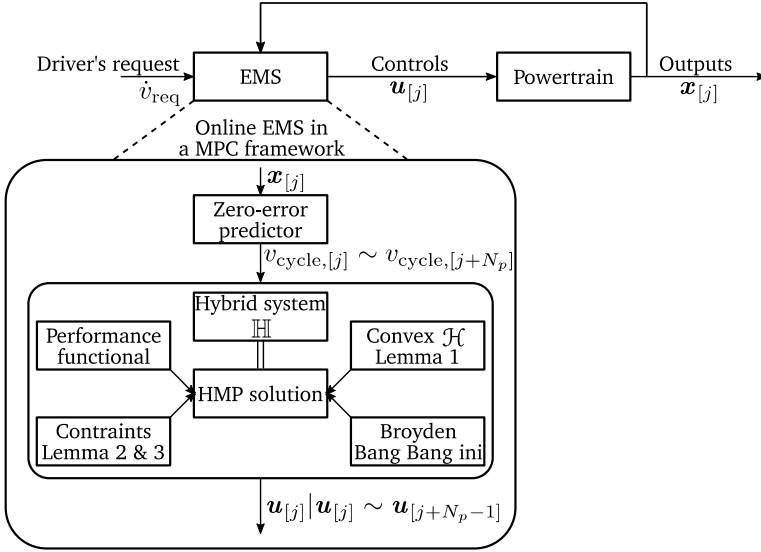


Figure 6.25: Structure of the predictive EMS with zero-error prediction.

The discrete state sequences of the EMS global optimal and the predictive EMS with the zero-error prediction are compared in Figure 6.26. The unnecessary shifting instances between 600 s and 1000 s that are noticed in the lower part of Figure 6.14 and Figure 6.20 are absent. At 767 s and 1153 s, both EMSs request the same gear position, which was not the case for the predictive EMS, as discussed in Section 6.3.3. With zero-error prediction, the predictive EMS reduces energy consumption in WLTC compared to the one with the MC predictor. It falls short by 0.01 kWh/100 km (0.10 %) compared to the EMS global optimal.

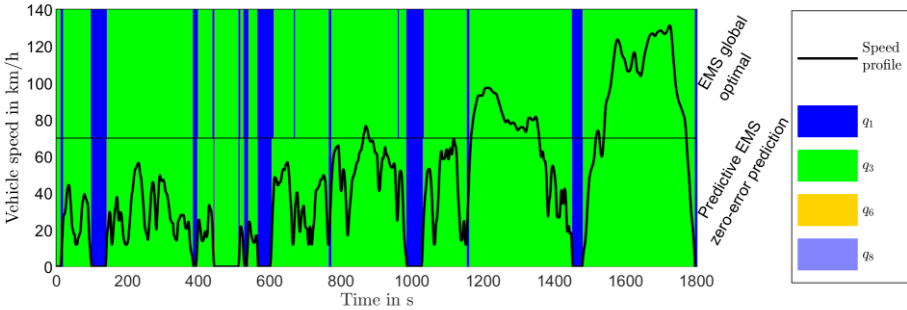


Figure 6.26: Discrete state sequences of the EMS global optimal and the predictive EMS with zero-error prediction

6.3.5 Influence of the Length of the Prediction Horizon

As mentioned in Chapter 3, the length of the prediction horizon N_p influences the performance of a predictive EMS [37]. Figure 6.27 shows the WLTC simulation results of the predictive EMS with N_p growing from 1 till 30. The energy consumption decreases from 11.540 kWh/100 km to 11.528 kWh/100 km, as the N_p increases from 1 to 10. At the same time, the shift count is reduced from 101 to 45. Rising N_p further reduces the energy consumption and the shift count until they saturate at the level of 11.525 kWh/100 km and 39, respectively. The predictive EMS with $N_p = 1$ is equivalent to minimizing the instantaneous energy consumption including the shift processes.

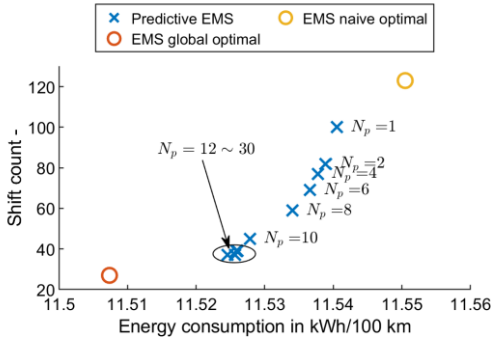


Figure 6.27: Energy consumptions and shift counts of the predictive EMS with different prediction horizon.

Figure 6.28 summarizes the elapsed time of the driving cycle simulations of the predictive EMS with different N_p . $N_p = 1$ is not included, since it requires no iteration. As analysed in Section 5.4,

the time complexity of **solveHMP** grows primarily linearly with the iteration number (N_{TPBVP}) and N_p . The elapsed time grows almost linearly between $N_p = 2$ and $N_p = 12$ in Figure 6.28, since their iteration numbers of each solution change little. Their average values are presented in Figure 6.29. Between $N_p = 2$ and $N_p = 12$, the average iteration numbers are close to 6. As N_p continues to increase, the elapsed time grows faster than linearly, since the average iteration number grows with N_p in the form of polynomials, as Figure 6.29 shows. Figure 6.30 summarizes the iteration number that each solution from **solveHMP** with $N_p = 30$ takes to converge. Averagely 10.2 iterations are required, which is marked by a red line. 917 solutions require 5 iterations to converge, while there are 1429 solutions in the case of the predictive EMS with $N_p = 10$, whose results are presented in Figure 6.13.

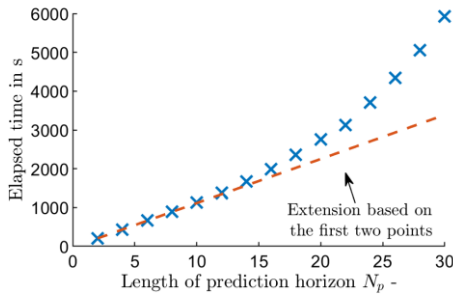


Figure 6.28: WLTC Simulation elapsed time⁵⁵ of the predictive EMS with different prediction horizon.

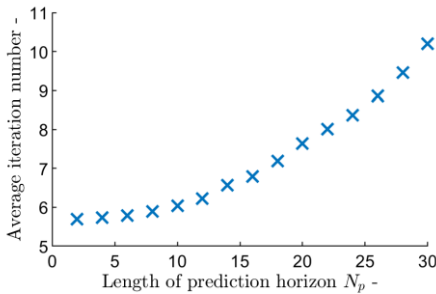


Figure 6.29: Average iteration numbers of the predictive EMS with different prediction horizon.

⁵⁵ The simulations are performed on a computer with Intel i5-8600K CPU with 3.60 GHz.

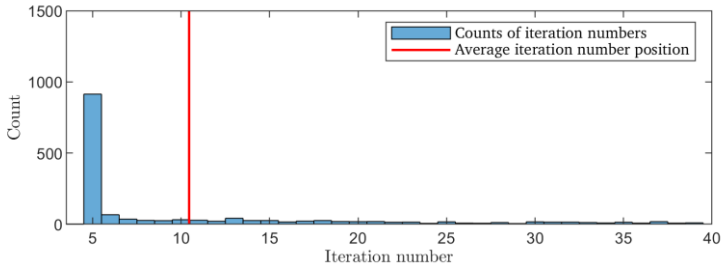


Figure 6.30: Iteration numbers for solutions in the predictive EMS to converge. $N_p = 30$.

With the saturation of the predictive EMS with $N_p = 12\sim 30$ in Figure 6.27 in mind, it is only necessary to consider a N_p smaller than 14. For a real-world application that requires the predictive EMS to respond at a fast pace, N_p is limited by the computational capacity of hardware.

6.4 Summary

In this chapter, the predictive EMS developed in Chapter 5 is investigated in a simulation environment from two perspectives: the solution process and the performance in the driving cycle simulation.

The solution processes for several representative driving conditions show how minimizing the instantaneous Hamiltonian leads to minimizing the cost functional. It also prevents consecutive shiftings between two gear positions that worsen the vehicle’s acceleration at high vehicle speed. The Broyden method with the Bang-Bang inspired initialization in `solveHMP` shows comparable convergence behaviour to the Pegasus method for a 1-D root finding example. It also shows obvious advantages, compared to applying the optimization tool box in Matlab to solve the 2-D TPBVP.

The energy consumption and shift count in the WLTC simulation of the predictive EMS is compared to those of the EMS naive optimal and the EMS global optimal, which are summarized in Figure 6.31. It is worth mentioning that the EMS naive optimal is a global optimal if shift processes of the vehicle model for evaluation are neglected, i.e. shifts happen instantly and shift processes consume or recuperate no energy. The metrics of the predictive EMS fall between the EMS naive optimal and the EMS global optimal. As the prediction horizon elongates, the performance of the predictive EMS approaches that of the EMS global optimal, until it meets a limit. Although not realistic, the limit can be exceeded by replacing the speed predictor to a zero-error predictor. Its performance represents the limit of pursuing a more accurate vehicle speed predictor for the purpose of predictive EMS.

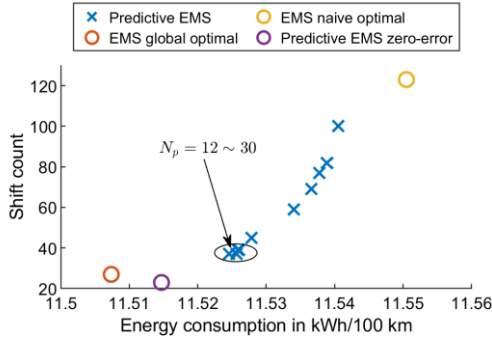


Figure 6.31: Energy consumptions and shift counts of all mentioned driving cycle simulation.

The elapsed time for simulation is analysed. Figure 6.32 presents its relationship to the multiplication of N_p and \bar{N}_{TPBVP} , the average iteration number of all solutions in a simulation. A supporting dashed line is generated by linear regression. It shows that the computation duration grows almost linearly with $N_p \cdot \bar{N}_{TPBVP}$, which provides evidence for the complexity analysis in Section 5.4. The nonlinearity is caused by the less dominant parts in the algorithm analysed in Section 5.4 as well as the different discrete state sequences in the simulations.

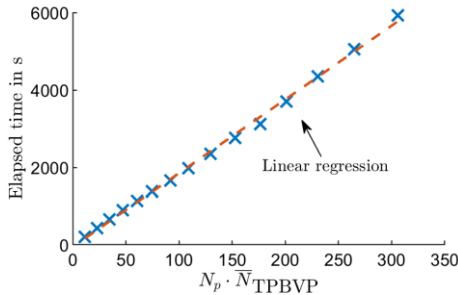


Figure 6.32: WLTC Simulation elapsed time of the predictive EMS.

It is observed that the EM1 propels primarily the vehicle through the WLTC. Furthermore, under the control of the predictive EMS, the ST2 stays away from the 2nd gear position. One reason is that Speed4E vehicle rarely drives with high speed and high acceleration under the WLTC. Figure 6.33 presents the discrete state sequences of the predictive EMS and the EMS naive optimal in a driving cycle simulation whose speed profile comes from the real-world driving data collected on a test vehicle of the IMS. The speed predictor is based on a Radials Function Neural Network chosen out of the comparative study [135]. Figure 6.32 compares their gear distributions. With high speed and high acceleration, both positive and negative, the predictive EMS requests the

vehicle to shift into the 2nd gear position. The predictive EMS reduces 84 % of the shift count (from 126, under the EMS naive optimal, to 20) in the driving cycle simulation. The energy consumption is reduced by 0.55 % (from 16.64 kWh/100 km to 16.55 kWh/100 km). However, it is still the EM1 that primarily propels the vehicle the ST2 in the neutral gear position. This information can be used to optimize the powertrain (selecting EMs and gear ratios).

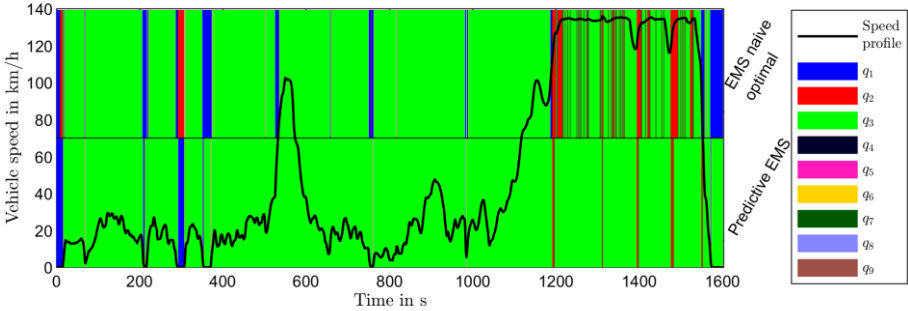


Figure 6.33: Discrete state sequences of the predictive EMS and the EMS naive optimal.

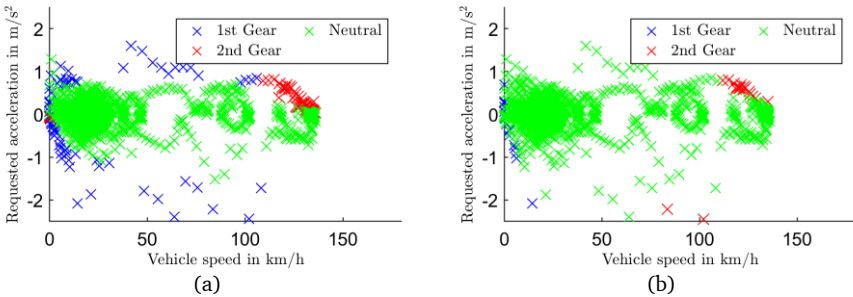


Figure 6.34: Gear distribution of (a) the EMS naive optimal and (b) the predictive EMS.

7 General Multi-Drive Multi-Speed Electric Vehicle

This chapter extends the results from Chapter 4 and Chapter 5 to propose a unified framework for designing EMSs for a general multi-drive multi-speed BEV with M EMs that each connect with a ST, $M \geq 2$, $M \in \mathbb{N}$. It can be imagined as the two-drive electric powertrain shown in Figure 3.1 with more EMs and STs parallelly coupled. However, as mentioned in Section 1.1, the general multi-drive multi-speed electric powertrain is meant for vehicles in a general sense, but not necessarily passenger cars. The continuous states of the system are the angular velocities $\boldsymbol{x} := (\omega_I, \omega_{II}, \dots, \omega_M)$. The continuous controls are the EM torques $\boldsymbol{u} := (T_I, T_{II}, \dots, T_M)$.

7.1 Hybrid System Formulation

The general multi-drive multi-speed BEV is modelled in this part with the assumption that all STs are either single-speed or share topology with the ST2 in Speed4E powertrain. Different classes of multi-drive electric powertrains can be achieved with different complexity. In the first two following subsections, all STs are multi-speed. Under such a setting, two different complexity levels can be realized through two assumptions. The first assumption for the lowest complexity: *one ST shifts at a time*. The second assumption for the highest complexity: *As many as possible STs shift at the same time*. The third subsection extends the model with the highest complexity. It is assumed that the number of multi-speed STs is smaller than M —the class that Speed4E powertrain belongs to.

7.1.1 One ST Shifts at a Time

It is assumed that *one ST shifts at a time*. Based on $\mathbb{H}_{\text{Speed4E}}$ in Section 4.4, the M -drive multi-speed powertrain can be represented by the hybrid automaton in Figure 7.1, in which M sub-automata help to model M EMs with their respective STs. Each sub-automaton takes over the model from Figure 4.5, whose switching conditions and sub-states are omitted for simplicity.

Importantly, the discrete state of the whole hybrid system is determined by a combination of the discrete states of all sub-automata, i.e.

$$\boldsymbol{\pi} := (q_I, q_{II}, \dots, q_M), \quad (7.1)$$

$$q_n \in \mathcal{Q}_n, \quad n \in \Pi, \quad (7.2)$$

$$\Pi := \{I, II, \dots, M\}, \quad (7.3)$$

$$\mathcal{Q}_n := \{q_{n,1}, q_{n,2}, \dots, q_{n,9}, q_{n,4,a}, q_{n,4,b}, \dots, q_{n,9,d}\}. \quad (7.4)$$

q_n is the discrete state of n -th sub-automaton, which is determined by ϖ_n , the discrete control in the n -th sub-automaton. The discrete set \mathcal{Q}_n includes the discrete states and the sub-states of the

sub-automaton that are defined in the same way as in Section 4.4. The discrete set of the whole hybrid system is $\mathcal{Q} := \{\mathcal{Q}_I, \mathcal{Q}_{II}, \dots, \mathcal{Q}_M\}$. For compactness, the discrete states and the sub-states representing the shift processes form a set $\mathcal{Q}_{\text{shift}}$. For the n -th sub-automaton, there is a set $\mathcal{Q}_{n,\text{shift}}$. Roman numerals are used for indexing to differentiate q_I and q_I in the hybrid system of Speed4E powertrain. They are also used for indexing entries in vectors.

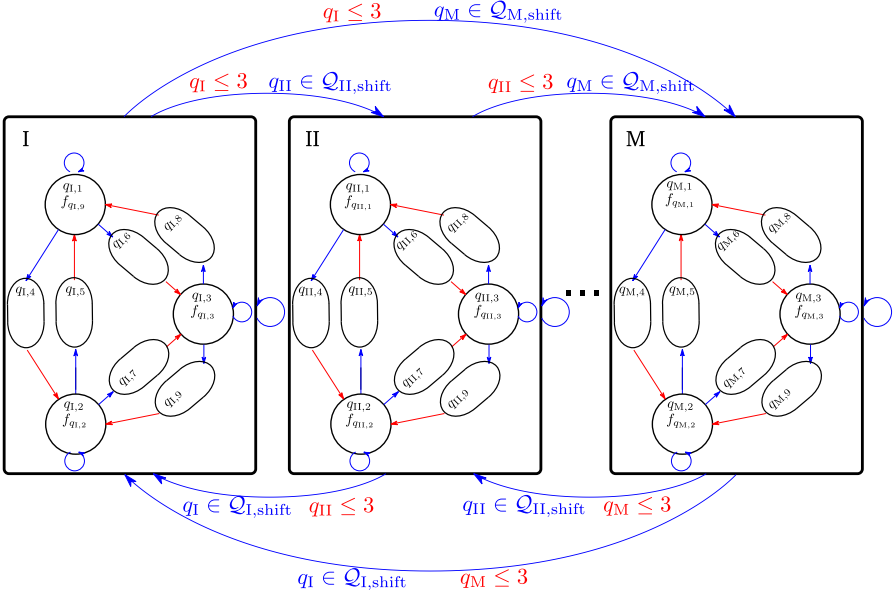


Figure 7.1: Hybrid automaton of an M-drive multi-speed powertrain with the first assumption.

Table 7.1: Admissible discrete control set, controlled switching.

| q_n | Admissible discrete control (\mathcal{B}_{q_n}) | | |
|--|---|---|---|
| 1 | 1 | 4 | 6 |
| 2 | 2 | 5 | 7 |
| 3 | 3 | 8 | 9 |
| $\forall n' \in \Pi, q_{n'}(t^-) \leq 3$ | | | |

Suppose that the n -th ST performs shifts, q_n changes to a discrete state representing shift processes. Given the assumption that one ST shifts at a time, the rest $q_{\Pi \setminus n}$ remains at their last values. The prerequisite is that all sub-automata are at their discrete states representing fixed gear positions. More specifically, at a time point t , a controlled switching can happen, if and only if $\forall n \in \Pi, q_n(t^-) \leq 3$. For intuitive understanding, a controlled switching can be seen as the activation of

a sub-automaton. If the last activated sub-automaton is at a discrete state representing a fixed gear position (the red conditions outside the sub-automata in Figure 7.1), a controlled switching can activate a new sub-automaton (the blue arrows outside the sub-automata in Figure 7.1). Table 7.1 defines the admissible control set for controlled switching of the n -th sub-automaton including the prerequisite for $\varpi_n(t) \in \{q_{n,4}, q_{n,5}, \dots, q_{n,9}\}$ in the red cell. The subscript n' is for differentiation from n in the table. Together with the admissible control set for autonomous switching from Table 4.3, the admissible discrete control set \mathcal{B}_n is formulated for the n -th sub-automaton. Note that \mathcal{B}_n evaluates the combination $\boldsymbol{\pi}$ for the prerequisite of the controlled switching.

At this stage, there can be multiple STs shifting at a time. To exclude this possibility, the sample space of $\boldsymbol{\pi}$ is defined as $\Omega_1 := \{(q_1, q_{II}, \dots, q_M) | \text{at most one } q_n \in \mathcal{Q}_{n,\text{shift}}, n \in \Pi, \Pi := \{I, II, \dots, M\}\}$. The discrete control ϖ_n is determined by $\boldsymbol{\pi}$ through \mathcal{B}_n . The discrete dynamics of the system is constrained by $\mathcal{B}_I, \mathcal{B}_{II}, \dots, \mathcal{B}_M$ and Ω_1 . The discrete transition function of the n -th sub-automaton is

$$q_n(t^+) = \varpi_n(t). \quad (7.5)$$

The dynamics of the hybrid system can be modelled similarly as in Section 4.4. The mathematical model is presented in Appendix C. The major difference is that the change of T_n during a shift process is compensated by the torques of all EMs excluding the n -th EM. The value of each torque remains to be determined by the EMS, which can be considered an optimization problem. Additionally, the reduced inertia of the powertrain depends on the discrete states of all sub-automata.

Similar to the HOCP in Section 4.5, the HOCP for the multi-drive multi-speed powertrain is

$$(\varpi_1^*(\cdot), \varpi_{II}^*(\cdot), \dots, \varpi_M^*(\cdot), \mathbf{u}^*(\cdot)) = \underset{\boldsymbol{\pi}(\cdot) \in \Omega_1, \varpi_n(\cdot) \in \mathcal{B}_{q_n(\cdot)}, \mathbf{u}(\cdot) \in \mathbf{U}(\cdot)}{\operatorname{argmin}} J(\boldsymbol{\pi}(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)), \quad (7.6)$$

$$J(\boldsymbol{\pi}(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)) = m(\mathbf{x}(t_f, p)) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t)) dt. \quad (7.7)$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + \sum_{n \in \Pi} PL_{EMn}(\omega_n, u_n), \quad (7.8)$$

$$m(\mathbf{x}(t_f)) = \beta \|\mathbf{x}(t_f) - \mathbf{x}_{t_f}\|^2, \quad (7.9)$$

$$\mathbf{x}_{t_f} = \frac{v_{\text{cycle}}(t_f)}{r} \cdot (i_1(q_I(t_f)), i_{II}(q_{II}(t_f)), \dots, i_M(q_M(t_f)))^T, \quad (7.10)$$

$$\text{subject to} \quad q_n(t_0) = q_{n,0}, \quad n \in \Pi, \quad (7.11)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad (7.12)$$

$$\mathbf{x}(t_f) = \mathbf{x}_{t_f}, \quad (7.13)$$

$$n \in \Pi, \quad (7.14)$$

$$u_n \in [T_{n,\min}(\omega_n), T_{n,\max}(\omega_n)], \quad n \in \Pi, \quad (7.15)$$

$$\omega_n \in [\omega_{n,\min}, \omega_{n,\max}], \quad n \in \Pi, \quad (7.16)$$

$$q_n \in \mathcal{Q}_n, \quad n \in \Pi, \quad (7.17)$$

$$\boldsymbol{\pi} := (q_I, q_{II}, \dots, q_M). \quad (7.18)$$

The cost functional is optimized over M number of discrete state sequences and a continuous control function. At a time point when the prerequisite of controlled switching is fulfilled, in total of $3M$ discrete instances can happen considering $n \in \Pi$ and $\omega_n \in \mathcal{B}_{q_n}$.

7.1.2 As Many as Possible STs Shift at the Same Time

The previous assumption can be loosened so that a M -drive multi-speed powertrain possesses more DoFs. It is assumed that *at most $M - 1$ STs can shift at the same time*. At least one EM propels the system during shift processes.

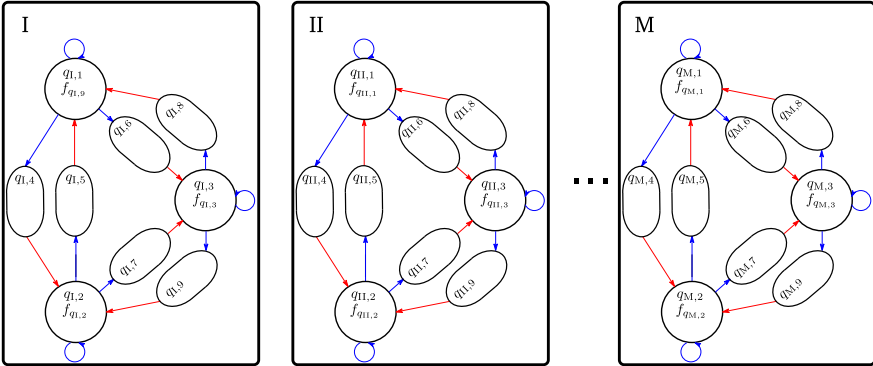


Figure 7.2: Hybrid automaton of a multi-drive multi-speed powertrain with the second assumption.

Figure 7.2 shows the hybrid automaton with the just-mentioned assumption. Activation and prerequisite conditions are no longer clear enough to be illustrated. All variables and conditions of Section 7.1.1 are taken over but the sample space. All sub-automata activate according to their admissible control sets $\mathcal{B}_I, \mathcal{B}_{II}, \dots, \mathcal{B}_M$ and a new sample space Ω_2 . The admissible control sets prevent any ST enters a shift process when other STs are shifting. To make sure that at most $M - 1$ STs shift at a time, the sample space of $\boldsymbol{\pi}$ is defined as $\Omega_2 := \{(q_I, q_{II}, \dots, q_M) | \text{at least one } q_n \in$

$$\mathcal{Q}_n \setminus \mathcal{Q}_{n,\text{shift}}, n \in \Pi, \Pi := \{I, II, \dots, M\}.$$

At a time point t , the indexes of the sub-automata in shift processes form a set Π_{shift} . Those remain in their fixed gear positions form a set $\Pi \setminus \Pi_{\text{shift}}$, which can not be an empty set according to Ω_2 . The change of the torques in the sub-automata from Π_{shift} is compensated by the torques in the sub-automata from $\Pi \setminus \Pi_{\text{shift}}$. The dynamics and the predefined control modelled for the system in Section 7.1.1 can be taken over. The necessary modification is presented in Appendix C.

The HOCP in Section 7.1.1 is taken over with the change of the sample space from Ω_1 to Ω_2 ,

$$(\varpi_I^*(\cdot), \varpi_{II}^*(\cdot), \dots, \varpi_M^*(\cdot), \mathbf{u}^*(\cdot)) = \underset{\pi(\cdot) \in \Omega_2, \varpi_n(\cdot) \in \mathcal{B}_{q_n(\cdot)}, \mathbf{u}(\cdot) \in \mathcal{U}(\cdot)}{\operatorname{argmin}} J(\boldsymbol{\pi}(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)). \quad (7.19)$$

In total of $M \cdot 3^{M-1}$ discrete instances can happen at a time point, when the prerequisite of controlled switching is fulfilled.

7.1.3 There are N multi-speed STs and N is smaller than M

Under this assumption, there are M EMs and N STs, $M \geq 2$, $1 \leq N < M$, $M, N \in \mathbb{N}$. The hybrid system is formulated based on the one in Section 7.1.2. In total of N STs are allowed to shift at the same time, for there is at least one EM connected with a single-speed ST propelling the vehicle. Figure 7.3 shows the hybrid automaton with the just-mentioned assumption.

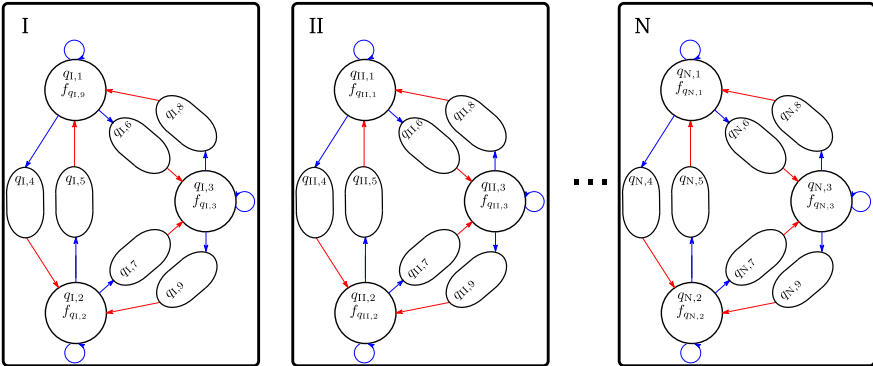


Figure 7.3: Hybrid automaton of a multi-drive multi-speed powertrain with the third assumption.

Since $N < M$, the combination $\boldsymbol{\pi}$ needs to be changed. $\boldsymbol{\pi}_3$ defines the discrete state of the whole system under the third assumption. It is defined as

$$\boldsymbol{\pi}_3 := (q_I, q_{II}, \dots, q_N), \quad (7.20)$$

$$q_n \in \mathcal{Q}_n, \quad n \in \Pi_N, \quad (7.21)$$

$$\Pi_M := \{I, II, \dots, M\}, \quad (7.22)$$

$$\Pi_N := \{I, II, \dots, N\}, \quad (7.23)$$

$$\mathcal{Q}_n := \{q_{n,1}, q_{n,2}, \dots, q_{n,9}, q_{n,4,a}, q_{n,4,b}, \dots, q_{n,9,d}\}. \quad (7.24)$$

The STs indexed by $N+I \sim M$ are single-speed. All variables of Section 7.1.2 are taken over but the sample space. Replace Π in Table 7.1 by Π_N and take over Table 4.3 to obtain the admissible discrete control sets. All sub-automata activate according to their admissible control sets $\mathcal{B}_I, \mathcal{B}_{II}, \dots, \mathcal{B}_N$ and a new sample space Ω_3 . The admissible control sets prevent any ST enters a shift process when other STs are shifting. At most N STs shift at a time, the sample space of $\boldsymbol{\pi}_3$ is defined as $\Omega_3 := \{(q_I, q_{II}, \dots, q_N)\}$.

At a time point t , the indexes of the sub-automata in shift processes form a set $\Pi_{N,\text{shift}}$. Those remain in their fixed gear positions form a set $\Pi_M \setminus \Pi_{N,\text{shift}}$, which can not be an empty set, since $N < M$. The change of the torques in the sub-automata from $\Pi_{N,\text{shift}}$ is compensated by the torques in the sub-automata from $\Pi_M \setminus \Pi_{N,\text{shift}}$. Take over the dynamics and the predefined control modelled for the system in Section 7.1.2 but replace Π_{shift} and $\Pi \setminus \Pi_{\text{shift}}$ with $\Pi_{N,\text{shift}}$ and $\Pi_M \setminus \Pi_{N,\text{shift}}$, respectively, to obtain the dynamics and the predefined control for the hybrid model with the third assumption.

Similar to the HOCB in Section 7.1.1 the general HOCB with the third assumption is

$$(\varpi_I^*(\cdot), \varpi_{II}^*(\cdot), \dots, \varpi_N^*(\cdot), \mathbf{u}^*(\cdot)) = \underset{\boldsymbol{\pi}_3(\cdot) \in \Omega_3, \varpi_{n'}(\cdot) \in \mathcal{B}_{q_{n'}(\cdot)}, \mathbf{u}(\cdot) \in \mathbf{U}(\cdot), n' \in \Pi_N}{\text{argmin}} J(\boldsymbol{\pi}_3(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)), \quad (7.25)$$

$$J(\boldsymbol{\pi}(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)) = m(\mathbf{x}(t_f, \mathbf{p})) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t)) dt. \quad (7.26)$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + \sum_{n \in \Pi_M} PL_{EMn}(\omega_n, u_n), \quad (7.27)$$

$$m(\mathbf{x}(t_f)) = \beta \|\mathbf{x}(t_f) - \mathbf{x}_{t_f}\|^2, \quad (7.28)$$

$$\mathbf{x}_{t_f} = \frac{v_{\text{cycle}}(t_f)}{r} \cdot (i_I(q_I(t_f)), i_{II}(q_{II}(t_f)), \dots, i_N(q_N(t_f)), i_{N+I}, \dots, i_M)^T, \quad (7.29)$$

$$\text{subject to} \quad q_n(t_0) = q_{n,0}, \quad n \in \Pi_M, \quad (7.30)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad (7.31)$$

$$\mathbf{x}(t_f) = \mathbf{x}_{t_f}, \quad (7.32)$$

$$n \in \Pi_M, \quad (7.33)$$

$$u_n \in [T_{n,\min}(\omega_n), T_{n,\max}(\omega_n)], \quad n \in \Pi_M, \quad (7.34)$$

$$\omega_n \in [\omega_{n,\min}, \omega_{n,\max}], \quad n \in \Pi_M, \quad (7.35)$$

$$q_n \in \mathcal{Q}_n, \quad n \in \Pi_M, \quad (7.36)$$

$$\boldsymbol{\pi}_3 := (q_I, q_{II}, \dots, q_N). \quad (7.37)$$

Note the fixed gear ratios of the STs indexed by $N+I \sim M$ in (7.29). The additional difference to the HOCPs (7.6) and (7.19) is that the discrete set in (7.25) is Π_N and in everywhere else is Π_M . In total of $N \cdot 3^N$ discrete instances can happen at a time point, when the prerequisite of controlled switching is fulfilled.

Set N to 1 and M to 2, the hybrid system represents Speed4E powertrain.

7.2 Hybrid Optimal Control of the General Multi-Drive Multi-Speed Electric Vehicle

The HOCP of Speed4E BEV is studied and a predictive EMS based on the HMP is developed for it. This subsection applies the HMP to the HOCPs of the general multi-drive multi-speed BEV under different assumptions whose hybrid systems formulated in Section 7.1. The HOCPs can be solved by the algorithms developed in Section 5.2 with optimality conditions produced in what follows.

7.2.1 One ST Shifts at a Time

In this subsection, all variables are defined according to Section 7.1.1.

As mentioned in Section 7.1.1, a combination $\boldsymbol{\pi} := (q_I, q_{II}, \dots, q_M)$ describes the discrete state of the whole system. Ω_1 is the sample space of $\boldsymbol{\pi}$. q_n , $n \in \Pi$, denotes the discrete state of a sub-automaton that represents an EM and a ST, and there are in total M STs. For the HOCP (7.6), a costates vector $\boldsymbol{\lambda} \in \mathbb{R}^M$, $\boldsymbol{\lambda} := (\lambda_I, \lambda_{II}, \dots, \lambda_M)^T$ is introduced to formulate the Hamiltonians. The discrete controls $\varpi_I, \varpi_{II}, \dots, \varpi_M$ change within their admissible sets $\mathcal{B}_I, \mathcal{B}_{II}, \dots, \mathcal{B}_M$, respectively, and $\boldsymbol{\pi}$ changes within Ω_1 . This causes The Hamiltonians, the costates, the continuous states and the continuous controls indexed by $\boldsymbol{\pi}$. The indexed Hamiltonians are defined as

$$\mathcal{H}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi, \boldsymbol{\lambda}_\pi) = l(\mathbf{x}_\pi, \mathbf{u}_\pi) + \boldsymbol{\lambda}_\pi^T \cdot \mathbf{f}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi). \quad (7.38)$$

Based on the hybrid model in Section 7.1.1 and Appendix C, The n -th entry of $\mathbf{f}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi)$ is

$$f_{\pi,n}(\mathbf{x}_\pi, \mathbf{u}_\pi) = \begin{cases} \frac{b(G, v, p(G, v, \hat{v}_{\text{pred}}))}{r} \cdot i_n(q_n) & \text{if } q_n \in \mathcal{Q}_n \setminus \{q_{n,j,c} \mid 3 < j \leq 9\}, \\ \frac{u_{\pi,n}}{\Theta_{\text{STn,input}}}, & \text{if } q_n \in \{q_{n,j,c} \mid 3 < j \leq 9\}. \end{cases} \quad (7.39)$$

$u_{\pi,n}$ denotes the n -th entry of \mathbf{u}_π . The dynamics of the indexed costates in (7.38) are

$$\dot{\lambda}_\pi(t) = -\frac{\partial \mathcal{H}_\pi}{\partial \mathbf{x}}(\mathbf{x}_\pi^*, \mathbf{u}_\pi^*, \lambda_\pi), \quad (7.40)$$

whose n -th entry is

$$\begin{aligned} & \dot{\lambda}_{\pi,n} \\ = & \begin{cases} -u_{\pi,n} - \frac{\partial PL_{EMn}}{\partial \omega_n}(\mathbf{x}_{\pi,n}, u_{\pi,n}) - \lambda_{\pi,n} \frac{\partial b}{\partial \omega_n}(G, v, T_{\text{total}}), & \text{if } q_n \in \mathcal{Q}_n \setminus \{q_{n,j,c} \mid 3 < j \leq 9\}, \\ -u_{\pi,n} - \frac{\partial PL_{EMn}}{\partial \omega_n}(\mathbf{x}_{\pi,n}, u_{\pi,n}), & \text{if } q_n \in \{q_{n,j,c} \mid 3 < j \leq 9\}. \end{cases} \end{aligned} \quad (7.41)$$

$\frac{\partial b}{\partial \omega_n}$ is fully expressed in (C.6).

Use an extended state $\mathbf{y}_\pi := (\mathbf{x}_\pi^\top, \boldsymbol{\lambda}_\pi^\top)^\top$ to keep the notation compact, whose dynamics follow

$$\dot{\mathbf{y}}_\pi = \Gamma_\pi(\mathbf{y}_\pi, \mathbf{u}_\pi^*) = \begin{pmatrix} \mathbf{f}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi) \\ -\frac{\partial \mathcal{H}_\pi}{\partial \mathbf{x}}(\mathbf{x}_\pi^*, \mathbf{u}_\pi^*, \lambda_\pi) \end{pmatrix}. \quad (7.42)$$

The hybrid controls consist of the combination $\varpi_1^*, \varpi_{II}^*, \dots, \varpi_M^*$ and \mathbf{u}^* . Similar to (5.23), (5.24) and (5.25), they are determined by

$$\mathbf{u}_\pi^* = \underset{\mathbf{u}_\pi \in \mathcal{U}}{\text{argmin}} \mathcal{H}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi, \lambda_\pi). \quad (7.43)$$

$$\varpi_1^*, \varpi_{II}^*, \dots, \varpi_M^* = \underset{\pi \in \Omega_1, \varpi_n \in \mathcal{B}_{q_n}, n \in \Pi}{\text{argmin}} \mathcal{H}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi^*, \lambda_\pi), \quad (7.44)$$

$$\mathbf{u}^* = \mathbf{u}_{\varpi_1^*, \varpi_{II}^*, \dots, \varpi_M^*}^*. \quad (7.45)$$

The transversality condition is

$$\lambda_{q(t_f)}(t_f) - 2\beta \left(\mathbf{x}(t_{f,p}) - \mathbf{x}_{t_{f,p}} \right) = \mathbf{0}. \quad (7.46)$$

The HOCP (7.6) can be solved by the algorithm developed in Section 5.2 with the conditions (7.38)-(7.46) provided by the HMP.

7.2.2 As Many as Possible STs Shift at the Same Time

For the HOCP (7.19), the HMP provides the same conditions as in Section 7.2.1 but replace (7.44) with

$$\varpi_1^*, \varpi_{11}^*, \dots, \varpi_M^* = \underset{\pi \in \Omega_2, \varpi_n \in \mathcal{B}_{q_n}, n \in \Pi}{\operatorname{argmin}} \mathcal{H}_\pi(\mathbf{x}_\pi, \mathbf{u}_\pi^*, \boldsymbol{\lambda}_\pi), \quad (7.47)$$

where the sample space is changed from Ω_1 to Ω_2 .

7.2.3 There are N multi-speed STs and N is smaller than M

The optimality conditions by the HMP for the HOCP (7.25) is largely same to those in Sections 7.2.1 and 7.2.2. The necessary modification focuses on the discrete sets Π_M and Π_N . A costates vector $\boldsymbol{\lambda} \in \mathbb{R}^M$, $\boldsymbol{\lambda} := (\lambda_1, \lambda_{11}, \dots, \lambda_M)^T$ is introduced to formulate the Hamiltonians. The discrete controls $\varpi_1, \varpi_{11}, \dots, \varpi_N$ change within their admissible sets $\mathcal{B}_1, \mathcal{B}_{11}, \dots, \mathcal{B}_N$, respectively, and $\boldsymbol{\pi}_3$ changes within Ω_3 . This causes The Hamiltonians, the costates, the continuous states and the continuous controls indexed by $\boldsymbol{\pi}_3$. The indexed Hamiltonians are defined as

$$\mathcal{H}_{\boldsymbol{\pi}_3}(\mathbf{x}_{\boldsymbol{\pi}_3}, \mathbf{u}_{\boldsymbol{\pi}_3}, \boldsymbol{\lambda}_{\boldsymbol{\pi}_3}) = l(\mathbf{x}_{\boldsymbol{\pi}_3}, \mathbf{u}_{\boldsymbol{\pi}_3}) + \boldsymbol{\lambda}_{\boldsymbol{\pi}_3}^T \cdot \mathbf{f}_{\boldsymbol{\pi}_3}(\mathbf{x}_{\boldsymbol{\pi}_3}, \mathbf{u}_{\boldsymbol{\pi}_3}). \quad (7.48)$$

Based on the hybrid model in Section 7.1.3, The n -th entry of $\mathbf{f}_{\boldsymbol{\pi}_3}(\mathbf{x}_{\boldsymbol{\pi}_3}, \mathbf{u}_{\boldsymbol{\pi}_3})$, $n \in \Pi_N$, is

$$\begin{aligned} & f_{\boldsymbol{\pi}_3, n}(\mathbf{x}_{\boldsymbol{\pi}_3}, \mathbf{u}_{\boldsymbol{\pi}_3}) \\ &= \begin{cases} \frac{b(G, v, p(G, v, \dot{v}_{\text{pred}}))}{r} \cdot i_n(q_n) & \text{if } q_n \in \mathcal{Q}_n \setminus \{q_{n,j,c} \mid 3 < j \leq 9\}, \\ \frac{u_{\boldsymbol{\pi}_3, n}}{\Theta_{\text{STn,input}}} & \text{if } q_n \in \{q_{n,j,c} \mid 3 < j \leq 9\}. \end{cases} \end{aligned} \quad (7.49)$$

$u_{\boldsymbol{\pi}_3, n}$ denotes the n -th entry of $\mathbf{u}_{\boldsymbol{\pi}_3}$. The n -th entry of $\mathbf{f}_{\boldsymbol{\pi}_3}(\mathbf{x}_{\boldsymbol{\pi}_3}, \mathbf{u}_{\boldsymbol{\pi}_3})$, $n \in \Pi_M \setminus \Pi_N$, is the same to the one with the condition $q_n \in \mathcal{Q}_n \setminus \{q_{n,j,c} \mid 3 < j \leq 9\}$ in (7.49).

The dynamics of the indexed costates in (7.48) are

$$\dot{\lambda}_{\pi_3}(t) = -\frac{\partial \mathcal{H}_{\pi_3}}{\partial \mathbf{x}}(\mathbf{x}_{\pi_3}^*, \mathbf{u}_{\pi_3}^*, \lambda_{\pi_3}), \quad (7.50)$$

whose π -th entry, $n \in \Pi_N$, $\dot{\lambda}_{\pi_3, n}$ equals

$$\begin{cases} -u_{\pi_3, n} - \frac{\partial PL_{EMn}}{\partial \omega_n}(x_{\pi_3, n}, u_{\pi_3, n}) - \lambda_{\pi_3, n} \frac{\partial b}{\partial \omega_n}(G, v, T_{\text{total}}), & \text{if } q_n \in \mathcal{Q}_n \setminus \{q_{n, j, c} | 3 < j \leq 9\}, \\ -u_{\pi_3, n} - \frac{\partial PL_{EMn}}{\partial \omega_n}(x_{\pi_3, n}, u_{\pi_3, n}), & \text{if } q_n \in \{q_{n, j, c} | 3 < j \leq 9\}. \end{cases} \quad (7.51)$$

$\frac{\partial b}{\partial \omega_n}$ is fully expressed in (C.6). The n -th entry of $\dot{\lambda}_{\pi_3}$, $n \in \Pi_M \setminus \Pi_N$, is the same to the one with the condition $q_n \in \mathcal{Q}_n \setminus \{q_{n, j, c} | 3 < j \leq 9\}$ in (7.51).

Use an extended state $\mathbf{y}_{\pi_3} := (\mathbf{x}_{\pi_3}^T, \lambda_{\pi_3}^T)^T$ to keep the notation compact, whose dynamics follow

$$\dot{\mathbf{y}}_{\pi_3} = \Gamma_{\pi_3}(\mathbf{y}_{\pi_3}, \mathbf{u}_{\pi_3}^*) = \begin{pmatrix} f_{\pi_3}(x_{\pi_3}, u_{\pi_3}) \\ -\frac{\partial \mathcal{H}_{\pi_3}}{\partial \mathbf{x}}(\mathbf{x}_{\pi_3}^*, \mathbf{u}_{\pi_3}^*, \lambda_{\pi_3}) \end{pmatrix}. \quad (7.52)$$

The hybrid controls consist of the combination $\omega_1^*, \omega_{11}^*, \dots, \omega_N^*$ and \mathbf{u}^* . Similar to (5.23), (5.24) and (5.25), they are determined by

$$\mathbf{u}_{\pi}^* = \underset{u_{\pi} \in \mathcal{U}}{\operatorname{argmin}} \mathcal{H}_{\pi}(x_{\pi}, u_{\pi}, \lambda_{\pi}). \quad (7.53)$$

$$\omega_1^*, \omega_{11}^*, \dots, \omega_M^* = \underset{\pi \in \Omega_3, \omega_n \in \mathcal{B}_{q_n}, n \in \Pi_N}{\operatorname{argmin}} \mathcal{H}_{\pi}(x_{\pi}, u_{\pi}^*, \lambda_{\pi}), \quad (7.54)$$

$$\mathbf{u}^* = \mathbf{u}_{\omega_1^*, \omega_{11}^*, \dots, \omega_N^*}^*. \quad (7.55)$$

The transversality condition is

$$\lambda_{q(t_f)}(t_f) - 2\beta(\mathbf{x}(t_{f,p}) - \mathbf{x}_{t_{f,p}}) = \mathbf{0}, \quad (7.56)$$

The HOCBP (7.6) can be solved by the algorithm developed in Section 5.2 with the conditions (7.48)-(7.56) provided by the HMP.

Set N to 1 and M to 2, the optimality conditions from this subsection is identical to those of Speed4E powertrain in Section 5.2.2.

7.3 Summary

A general multi-drive multi-speed electric powertrain is proposed. Its hybrid system formulation is presented. In the formulation, each EM is connected to a ST that is either single-speed or shares the same topology with the ST2 in Speed4E powertrain. By doing so, the methodology to model the shift processes is taken over. The mathematical models of the multi-drive multi-speed powertrain under two assumptions are separately presented. It is noticed that they share the majority of their hybrid system formulation. The lowest and the highest complexity caused by the different assumptions is expressed through the admissible discrete control sets and the sample space of the combination that defines the discrete state of the whole system. In the same way, the general multi-drive multi-speed electric powertrain with other complexity can be modelled. The HOCPs for the multi-drive multi-speed powertrain are presented. They are in nature under the same category as the HOCP for Speed4E powertrain.

The multi-drive multi-speed powertrain in Section 7.1 does not cover the powertrains that consist of STs with different topologies to the ST2 of Speed4E powertrain. Though the difference is seemingly giant, the powertrains that are not explicitly included in Section 7.1 can be modelled in the same way as in Section 7.1, as long as the shift processes are modelled with the same methodology as in 4.4.2, i.e. simplifying the shift processes to be primarily time dependent. On the level of hybrid automata, the sub-automata in Figure 7.1 and Figure 7.2 are to be modified, but the switching conditions among sub-automata remain. The HOCPs in Section 7.1 remain.

Lemma 1, 2 and 3 from Section 5.3 can be extended to higher dimensional problems obtained in Section 7.2. The corresponding problem is formulated as

$$\min_{\mathbf{u}_\pi} l(\mathbf{x}, \mathbf{u}_\pi), \quad (7.57)$$

$$l(\mathbf{x}, \mathbf{u}_\pi) = \mathbf{x}^T \cdot \mathbf{u}_\pi + \sum_{n \in \Pi} PL_{EMn}(\omega_n, u_n), \quad (7.58)$$

$$\text{subject to} \quad \dot{\mathbf{x}}_\pi = \frac{b(q, v, T_{\text{total}})}{r} \cdot \left(i_I(q_I(t_f)), i_{II}(q_{II}(t_f)), \dots, i_M(q_M(t_f)) \right)^T, \quad (7.59)$$

$$T_{\text{total}} = \sum_{n \in \Pi} \phi_n(\omega_n, u_n) \cdot i_n(q_n), \quad (7.60)$$

$$\Pi := \{I, II, \dots, M\}, \quad (7.61)$$

$$u_n \in [T_{n,\min}(\omega_n), T_{n,\max}(\omega_n)], \quad n \in \Pi, \quad (7.62)$$

$$\omega_n \in [\omega_{n,\min}, \omega_{n,\max}], \quad n \in \Pi, \quad (7.63)$$

$$\boldsymbol{\pi} := (q_I, q_{II}, \dots, q_M), \quad (7.64)$$

$$q_n \in \{1, 2\}, \quad n \in \Pi. \quad (*)$$

The HMP is applied to the HOCPs of the general multi-drive multi-speed BEV from Section 7.1. The difference between the optimality conditions of both HOCPs is the sample space of the combination. A unified framework for modelling and optimal control of the general multi-drive multi-speed electric powertrain takes shape based on the hybrid system formulation and the HOCP formulation in Section 7.1 as well as the optimality conditions from the HMP in Section 7.2.

The predictive EMS developed in Chapter 5 and both reference EMSs in Chapter 6 can be applied. Compared to the EMS naive optimal, it can be expected that the predictive EMS reduces the shift count and the energy consumption of a multi-drive multi-speed BEV, since the advantages of the predictive EMS over the EMS naive optimal is not Speed4E powertrain specific.

Due to the lack of data, the HOCP of the general multi-drive multi-speed BEV is not numerically studied.

8 Multi-Criteria Operation Strategy Considering Service Life

As Figure 6.18 shows, the vehicle is propelled mostly by the EM1 through the ST1 in the WLTC, may it be controlled by the predictive EMS, the EMS naive optimal or the EMS global optimal. One can imagine that the EM1 and the ST1 would reach their service lives earlier than the EM2 and the ST2, if they would be designed to reach the same service life under the WLTC. Such a hypothetical case leads to two possible outcomes: 1) if certain parts in the EM1 or the ST1 that reach their service lives are not independently repairable or replaceable, the EM1 or the ST1 is discarded; 2) if they are independently repaired or replaced, maintenance and re-calibration costs occur. Both outcomes are undesirable, which are addressed often during design processes [36, p. 279]. This chapter proposes a solution to avoid such a hypothetical case from an operation strategy point of view. With such a goal in mind, this chapter does not pursue innovative or comprehensive methods but conventional and proved in practice methods to characterize service life.

Firstly, the relevant fundamentals surrounding the concept “service life” are introduced and discussed, which leads to a quantitative service life estimation model of Speed4E powertrain. Secondly, with the help of the service life estimation, the goal of avoiding the hypothetical case in the previous paragraph is translated to mathematical language and a multi-criteria operation strategy (MCOS) is formed. Thereafter, the influence of the controls on the service life is investigated. Lastly, the driving cycle simulation results of the developed MCOS are discussed.

8.1 Service Life Estimation

8.1.1 Introduction

Service life is more precisely defined as B_x -*service life*, which indicates a service life B , within which x % of the products of a production batch fail [36, p. 59]. It may have a unit of duration, e.g. the service life of an electric motor in hour, or a unit of distance, e.g. the service life of a transmission in km. In some literatures, they are called *lifetime* and B_x -*lifetime*, respectively [153]. B_x -service life describes the failure and reliability characteristics of a product. Two core ideas are: *Reliability* is the probability that a product does not fail during a defined period of time under given functional and surrounding conditions [153, p. 21]; *Failure rate* is the complement of the reliability.

Figure 8.1 qualitatively illustrates a typical bathtub curve of failure rates along service life. The bathtub curve can be divided into three regions: region 1 of early failures, region 2 of random failures and region 3 of wearout failures. In Region 1, the failure rate decreases, as the service life increases. Early failures are mostly caused by errors in assembly, production, materials or by design flaws. The failure rate remains constant in Region 2. Often, incorrect operation, maintenance and

dirt particles are the main causes for random failures. They are difficult to predict and happen with a low probability. The failure rate increases rapidly, as the service life increases and enters Region 3. As mechanical parts fatigue and decayable materials age, a wearout failure takes place with an increasing probability.

The actions taken to reduce failure rates of early failures and random failures must be taken into consideration in design processes. From a system design perspective, the improvements in Region 3 take place at the stage of construction dimensioning [36, p. 35]. Wearout failures are most interesting for design and reliability analysis, for B_x -service life is largely determined by their failure rates. They can be reduced substantially by taking special account of possible causes and designing the parts and the systems accordingly, which leads to service life calculation [153, p. 635].

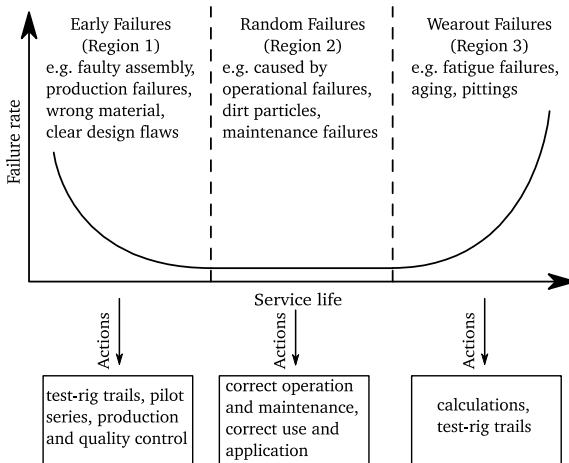


Figure 8.1: The bathtub curve

The service life of a product depends on the service lives of its individual parts and their collaboration. For instance, a passenger car with a 5-speed transmission drives on a given driving cycle, during which different gears are engaged for different time. Figure 8.2 shows the proportion of time of all gears. To calculate the service life of the transmission in the example is a combined consideration of the service lives of the parts corresponding the power flow of each gear position.

Other than time, loads and, hence, stresses on a part is also crucial to the service life calculation. They depend on its functionals and the surrounding conditions. For the transmission in the example, the loads and stresses on each part depend on the driving cycle. Considered over time, indi-

vidually variable load curves of all parts are obtained. They are compared against the parts' resistance to loads and stresses to obtain accumulated fatigue and service lives [36, p. 251].

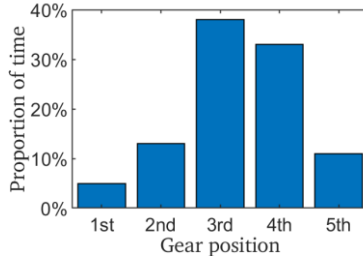


Figure 8.2: Proportion of time that different gears are engaged, when the vehicle drives with an example driving cycle.

8.1.2 Parts Selection for Service Life Estimation

For electric powertrains, EMs and transmissions are prone to wearout failures. In order narrow down the subjects for service life estimation, their parts are analysed and divided into three categories, as shown in Table 8.1. A-parts are defined as critical to reliability and their service lives can be calculated corresponding reality. B-parts are defined as critical to reliability, but the service life is hard to be calculated with confidence. Often, these parts are loaded by friction, abrasion (synchronizers and brushes), dirt and corrosion (seals). C-parts are neutral to risk and not considered in further discussion. Service life estimation focuses on A-parts.

Table 8.1 Preselection of parts for fatigue estimation

| | A-parts Prone to risk B_x -service life calculation possible | B-parts Prone to risk B_x - service life calculation impossible or does not reflect reality | C-parts Neutral to risk Mostly random and early failures |
|---|---|--|---|
| Electric motor [154], [155] | Winding insulation Bearing | Brush (if exists) Rotor and end ring (asynchronous motors) | |
| Transmission [153, p. 99], [36, p. 647] | Shaft Gear Bearing | Synchronizer Seal | Housing |

In the category “A-parts”, winding insulation is different from the rest for its main purpose. It protects winding from electrically shorting, contamination, mechanical and thermal stress [156, pp. 7-12]. The electrical endurance of insulation materials is affected by temperature and time in

such a way that obeys Arrhenius chemical rate equation, i.e. the thermal service life⁵⁶ of insulation is halved for each increase of 10 °C⁵⁷ in the exposure temperature [158, 159]. It is natural and essential, in the case of any electric powertrain, to take the cooling system of the electric motor into the consideration. The mathematical modelling of the complex thermal management system developed for Speed4E powertrain is, unfortunately, beyond the scope of online MCOS in this work. However, another work of the author [160] showed that a goal-oriented thermal management control can not only reduce the energy consumption but also extend the thermal service lives of both EMs in Speed4E powertrain.

The rest of the A-parts are designed either for operational fatigue strength or for endurance strength [36, pp. 294-296]. To design a part for operational fatigue strength is to have it designed to endure an expected load profile for a certain time of use [36, p. 250]. To design a part for endurance strength is to base the design on the maximum anticipated load [36, p. 279].

Shafts are often conservatively designed for endurance strength, for the typical shaft configuration of vehicle transmissions is particularly unfavourable from the point of view of strength. The long distance between bearings causes large bending moments, and the contacting parts such as shoulders, grooves, collars, bearing seats etc. cause notches [36, p. 280]. As a result of conservative designs, shafts in transmissions are often found fatigue resistant, i.e. infinite service life, in service life estimation for normal expected load profile, as shown in a typical example in Table 8.2. Therefore, shafts are not considered in the service life estimation in this work.

Table 8.2. An typical example of calculated B_1 and B_{10} service life of A-parts in a transmission [153, p. 102]

| Components | Estimated service life |
|---|--|
| Input shaft | Fatigue resistant |
| Output shaft | Fatigue resistant |
| Gear tooth failure: first stage gearset | B_1 service life: 70,000 to 120,000 cycles |
| Gear tooth pitting: first stage gearset | B_1 service life: 500,000 cycles |
| Rolling bearing | B_{10} service life: 1,500,000 cycles to fatigue resistant |

The B_1 service lives of the first stage gearset for tooth failure and pitting are listed in Table 8.2.

⁵⁶ In the literature [157], thermal service life is called thermal life and thermal lifetime.

⁵⁷ In actual testing, the increment often deviates from 10°C.

There are different types of failures for a part, and their failure rates and B_x -service lives are different. Rolling bearings, depending on the configuration, can operate with a B_{10} service life varying from 1,500,000 cycles to an infinite service life. Typical service lives of gears and bearings that are reasonably designed for a system are not decidedly separated, which suggests that both types of parts are necessary to be considered in service life estimation.

In summary, gears and bearings are selected as the subjects of the service life estimation. During operation, the fatigue of these parts accumulates, as they withstand loads. The following elaborates, how the fatigue and its accumulation are modelled.

8.1.3 Mechanical Fatigue Calculation and Its Accumulation

Gears and bearings are exposed to different types of failures, among which some are caused by fatigue and under the category of wearout failures, while others are not.

8.1.3.1 Main Damages of Gears and Rolling Bearings

Gear

The service life of a gear is basically limited by four types of damages [36, pp. 239-243]: 1) tooth failure, 2) macropitting and micropitting, 3) hot scuffing and 4) cold scuffing. Both tooth failure and macropitting can lead to gear failure. Micropitting is surface pitting in a microscopic scale. It does not lead to gear failure but can increase noise emission. Hot scuffing and cold scuffing are related to lubrication condition and seldom occur in vehicle transmissions. They do not cause gear failure. For their seriousness and relevancy, tooth failure and macropitting are necessary to be considered in the service life estimation. Read [36, pp. 239-243] for more information.

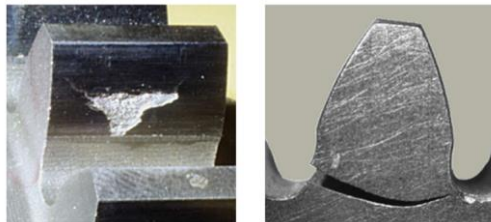


Figure 8.3: Typical damage examples. (left) Macropitting. (right) Tooth root breakage—tooth failure [161]

Tooth failure is that the whole tooth or part of a tooth breaks off, which leads to gear failure. Figure 8.3 (right) shows an example of tooth root breakage, which will cause the whole tooth to break off. The tooth operates under pulsating load and the maximum bending stress occurs at the tooth root, which suggests that the fatigue at the tooth root, i.e. tooth root breakage, needs to be

modelled to represent tooth failure [36, p. 242].

Damage to the tooth flank by pitting is indicated by the appearance of pin holes and extended tooth flank spalling, which is a symptom of material fatigue at the tooth flanks. See Figure 8.3 (left) for an example. It is considered mostly caused by the stresses between the flanks of the pinion and the wheel during meshing. The fatigue on the tooth flanks is to be modelled.

Rolling Bearing

Loads on bearings in transmissions result primarily from the meshing forces of the gears mounted on the same shafts and preload forces with adjusted bearing arrangements [36, p. 423]. Rolling bearing damages can be divided into following categories: adhesive wear, abrasive wear, fatigue, corrosion, and further failures (electric erosion, plastic deformation etc.) [162]. Only fatigue is considered in the service life estimation, since adhesive wear, abrasive wear and corrosion are mainly caused by false installation, operation, and maintenance; methods to estimate the failures caused by electric erosion, plastic deformation etc. are out of the scope of the service life estimation for online MCOSs [163].



Figure 8.4: Spalling on the inner ring of a rolling bearing [167].

Fatigue occurs at surfaces of rolling elements and rings, due to cyclic stresses from rolling and sliding between rolling elements and rings. They cause microstructural alterations and lead to pitting and spalling on surfaces [164]. Figure 8.4 shows an example of the spalling on the inner ring caused by fatigue.

Fatigue to be modelled in the service life estimation

So far, tooth root breakage and macropitting (it is called “pitting” in what follows) are selected as the failures considered in the service life estimation for gears. It was stated in [36, p. 239] that *gears for vehicle transmissions are now almost without exception designed so that tooth root breakage is eliminated, for it causes serious and immediate gear failure*. It will be evaluated in Section 8.1.4 to decide, if tooth root breakage remains in the service life estimation. Bearing fatigue is selected as the failure considered in the service life estimation for bearings. In a vehicle, gears and bearings

withstand loads caused by both positive and negative torques generated by its power source(s). For conventional vehicles, [165] and [166, pp. 110-111] suggested that loads and fatigue caused by negative torques are low enough to be neglectable in the service life estimation. Such a practice, however, loses its prerequisite in the case of BEVs, since they primarily use their EM(s) to recuperate their kinetic energy while decelerating. In this case, loads and fatigue caused by positive and negative torques are comparable, which is verified later in Section 8.1.4. It is necessary to discuss the available methods to model fatigue caused by both positive and negative torques.

- **Tooth root breakage**

Bending stress is considered the only stress that is critical for modelling tooth root breakage [36, p. 243]. Bending stress on a tooth root is schematically illustrated in Figure 8.5. The bending stress caused by the normal force on the right flank ($F_{n,right}$) is marked blue and the bending stress caused by $F_{n,left}$ is marked red, both of which contribute to the tooth root breakage. It was suggested in [161] to consider the alternating loads on both flanks and the residual stresses in the tooth core to model the fatigue of the tooth root breakage. Additionally, DIN 3990—part 6 suggested four methods:

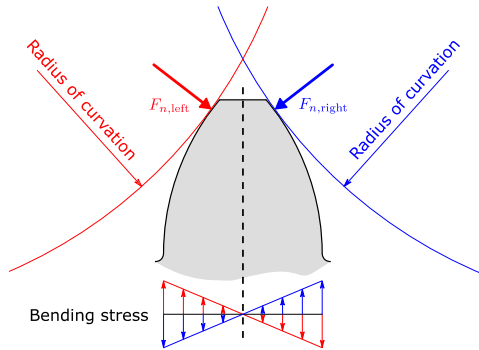


Figure 8.5: Schematic of the bending stress on the tooth root.

- *Method FA*: the fatigue is estimated with a damage characteristic curve, i.e. a S-N curve discussed in next subsection, based on calculation or experiment that considers the frequency of the alternating loads.
- *Method FB*: the fatigue is estimated with a S-N curve that is modified considering the frequency of the alternating loads.
- *Method FC*: the loads caused by negative torques are added to the loads caused by positive torques with a factor of 0.7. The combined loads are used to estimate the total fatigue based on a S-N curve.

- *Method FD*: the loads caused by negative torques are multiplied with 0.7 and the resulting loads are used to estimate a partial fatigue based on a S-N curve. The partial fatigue is added to the fatigue that is estimated with the loads caused by negative torques based on a S-N curve.

The method from [161] and the methods FA and FB are not applicable, since the complexity exceeds the scope of the service life estimation for MCOSs. The methods FC and FD were mostly applied to conventional vehicles, which decelerates majorly through brakes, while gears in transmissions only transmit small negative torques generated by engines. In this case, there is a decided distinction between the loads caused by the positive and the negative torques, which is not the case for BEVs. For the purpose of conservative estimation, the method FC is adopted but with a conservative factor of 1.0 for the loads caused by negative torques.

- Tooth flank pitting

Pitting is caused by slip/roll stresses on flank surfaces and incipient cracks due to high shear stresses in the area below surfaces [36, p. 239]. The schematic of the stresses on the surface is shown in Figure 8.6. It is assumed none interference between the slip/roll stresses on both flanks, as well as between the shear stresses below both flanks. The loads caused by positive torques and negative torques are, therefore, separately calculated and the resulting fatigues are separately accumulated.

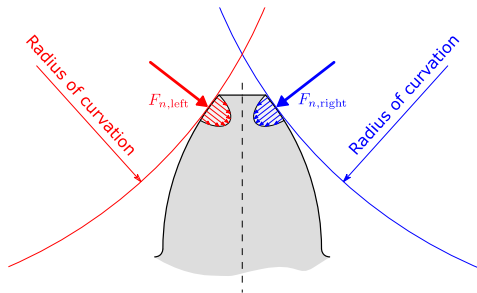


Figure 8.6: Schematic of the Hertzian stress on the tooth flanks

- Bearing fatigue failure

The theory for assessing material deformation in rolling bearings is adopted from [167], which assumes that cracks are caused by the maximum orthogonal shear stresses initiated in the risk volume illustrated in Figure 8.7. Under the assumption from [167], for a standard bearing configuration, i.e. two bearings mounted on a shaft, the assessment of stresses, risk volume and resulting fatigue is simplified to be free from the directions of torques. The loads

caused by positive and negative torques are linearly accumulated.

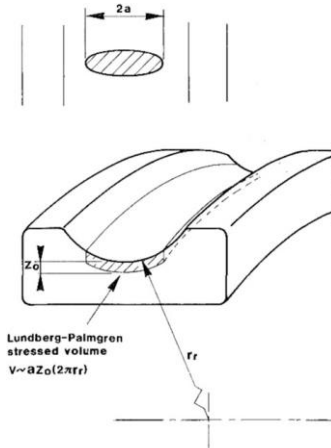


Figure 8.7: Risk volume of rolling bearing fatigue suggested by [113]. Figure taken from [114].

After analysis, tooth flank pitting and rolling bearing fatigue are chosen to be the subjects of the service life estimation. The fatigues caused by positive and negative torques on the former are separately accumulated and those on the latter are linearly accumulated.

8.1.3.2 Fatigue Estimation

To estimate the fatigue resulting in a certain kind of failure is to compare the load causing the fatigue to the resistance of the load, i.e. load capacity. It is standard to be modelled as a S-N curve, i.e. a Stress-Life Cycle curve. It is a stress-controlled Wöhler curve, which describes the load capacity of a component. It is attained through experiments⁵⁸ to reflect the relationship between the cyclic stress amplitude and the permissible cycle number under the cyclic stress [168, pp. 21-25]. Figure 8.8 shows the typical shape of S-N curves on double logarithmic scales. The curves of different failure rates are marked with different colours, among which the one of B_{10} -service life is

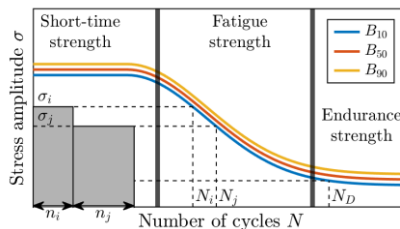


Figure 8.8: S-N curves of different B_x -service life of an example part for a certain failure. Figure modified from the one from [32, p. 254].

mostly used for gears and bearings. The permissible cycle number N_i of B_x -service life can be read from a curve for a given stress amplitude σ_i . Three zones can be identified:

- *Short-time strength*, permissible cycles up to ca. $10^1 \sim 10^3$ cycles: also called *low cycle fatigue* domain. Fatigue in this zone is caused by irreversible plastic deformations in materials due to high stress [153, p. 321, 169].
- *Fatigue strength*, permissible cycles up to the cycle number at fatigue limit $N_D = 10^6 \sim 10^7$: also called *high cycle fatigue* domain. In this zone, the S-N curve can be simplified as a straight line in double logarithmic coordinates, as

$$N_i = N_D \cdot \left(\frac{\sigma_i}{\sigma_D} \right)^{-k}, \quad (8.1)$$

where σ_D is the fatigue limit stress amplitude and k is the gradient of the straight line in double logarithmic coordinates [36, p. 254].

- *Endurance strength*, the stress amplitude is smaller than σ_D : no failure is expected in this zone for a wide range of materials, e.g. ferritic steels. Some materials, e.g. austenitic steels, do not show a distinct endurance strength [153, p. 321].

A part is ideally designed in such a way that the expected operational stresses do not fall into the short-time strength zone and avoid the endurance strength zone, avoiding short expected service life and over dimensional design, respectively. S-N curves are often simplified to be horizontal on double logarithmic scales in the short-time strength and the endurance strength zones.

- Tooth root breakage

DIN 3990—part 3 provides the standard method to calculate S-N curves of B_{10} service life of tooth root breakage. See [36, pp. 243-245] for a detailed explanation. See Appendix E for the calculated S-N curves in this thesis.

The effective bending stress σ_F characterizes the local stress of a tooth root. According to DIN 3990, σ_F can be determined by considering F_{tn} , the tangential component of the normal force, and various parameters,

$$\sigma_F = \frac{F_{tn}}{b \cdot m_n} \cdot Z_{Fa} \cdot Z_{Sa} \cdot Z_\epsilon \cdot Z_\beta \cdot K_A \cdot K_V \cdot K_\alpha \cdot K_\beta, \quad (8.2)$$

where the parameters are defined in Table 8.3.

- Tooth flank pitting

DIN 3990—part 2 provides the standard method to calculate S-N curves of B_{10} service life

off tooth flank pitting. See [36, pp. 245-246] for a detailed explanation. See Appendix E for the calculated S-N curves in this thesis.

Hertzian stress σ_H characterizes slip/roll stresses on flank surfaces According to DIN 3990, it can be determined by considering the nominal force on the tooth flank F_n and various parameters,

$$\sigma_H = Z_{B/D} \cdot Z_H \cdot Z_E \cdot Z_\epsilon \cdot Z_\beta \cdot \sqrt{\frac{F_n(i+1)}{d_1 \cdot b \cdot i}} \cdot K_A \cdot K_V \cdot K_\alpha \cdot K_\beta, \quad (8.3)$$

where the parameters are defined in Table 8.3.

Table 8.3. Parameters to calculate σ_H

| Parameter | Meaning | Source |
|--------------|--|---|
| $Z_{B/D}$ | pinion/wheel contact factor | DIN 3990—Part 2, Pages 8 |
| Z_H | zone factor | DIN 3990—Part 2, Page 6 |
| Z_E | elasticity factor | DIN 3990—Part 2, Page 8 |
| Z_ϵ | contact ratio | DIN 3990—Part 3 for σ_F and Part 2 for σ_H |
| Z_β | helical overlap | DIN 3990— Part 3 for σ_F and Part 2 for σ_H |
| d_1 | Pitch circle diameter of the pinion | gears design |
| b | contact face width | gears design |
| m_n | standard module | gears design |
| i | gear ratio | gears design |
| K_A | application factor | DIN 3990—Part 1, Page 55 |
| K_V | dynamic factor | DIN 3990—Part 1, Page 16-17 |
| K_α | transverse factor | DIN 3990—Part 1, Page 45 |
| K_β | longitudinal load distribution factor for surface stress | DIN 3990—Part 1, Page 19 |

- Bearing fatigue failure

DIN ISO 281 provides the standard method to calculate B_{10} service lives of rolling bearings directly. N , the permissible cycle number under an equivalent dynamic load L , is

$$N = \left(\frac{C}{L}\right)^k, \quad (8.4)$$

$$L = X \cdot F_{\text{radial}} + Y \cdot F_{\text{axial}}, \quad (8.5)$$

where the variables and the parameters are defined in Table 8.4.

Table 8.4. Parameters to calculate N

| Parameter | Meaning | Source |
|---------------------|--------------------------------|--|
| C | basic dynamic load rating in N | Bearing catalogues |
| L | equivalent dynamic load in N | calculation |
| k | service life exponent | DIN ISO 281. Bearing type dependent |
| X | radial factor | DIN ISO 281. F_r and F_a dependent |
| F_{radial} | radial load in N | calculation |
| Y | axial factor | DIN ISO 281. F_r and F_a dependent |
| F_{axial} | axial load in N | calculation |

8.1.3.3 Fatigue Accumulation

The fatigue accumulation hypothesis describes, how the fatigues caused by cyclic stresses integrate, whose foundation is Miner's rule [170]. It assumes that the absorbed work during operation causes fatigue and the proportion of the already absorbed work w out of the permissible work to absorb W measures the fatigue. Its proportion is equal to the one of the cycle number of a cyclic stresses with constant amplitude n to the permissible cycle number N . Under such an assumption⁵⁹, every load cycle causes a constant fatigue, which accumulates linearly. Furthermore, the

⁵⁹ Methods with the assumption of nonlinear accumulation showed high accuracy (see the review article [171]). However, they were targeted at specific use cases and demand high computation resources [166, p. 26], which makes them unfavourable for the service life estimation in MCOSs.

accumulated fatigue D caused by cyclic stresses with j amount of stress amplitudes, i.e. a load profile, can be summed up as

$$D = \frac{w_1}{W} + \frac{w_2}{W} + \dots + \frac{w_j}{W} = \frac{n_1}{N_1} + \frac{n_2}{N_2} + \dots + \frac{n_j}{N_j} = \sum_{i=1}^j \frac{n_i}{N_i}, \quad (8.6)$$

where i indexes the stress amplitude. The part in Figure 8.8 has operated under σ_i and σ_j for n_i and n_j cycles, respectively. D is the sum of the fractions of the areas with the width of n_i and n_j out of the areas with the width of N_i and N_j . In continuous time, D can be determined by

$$D = \int_{t_0}^{t_f} \frac{n(t)}{N(\sigma(t))} dt. \quad (8.7)$$

A failure is expected, when D reaches 1 [170, 172]. In reality, a failure often occurs with a calculated fatigue smaller than 1. In order to use (8.6) to perform an accurate prediction, the knowledge of the total fatigue, when an actual failure happens, is necessary [36, p. 258].

With a simplified S-N curve that is horizontal in endurance strength zone, the fatigues caused by the stresses under the fatigue limit stress can be ignored, since an infinite service life would be expected. However, it contradicts the results reported in [173, 174], which stated that, in a mixed load profile, stresses under the fatigue limit stress cause fatigue. Therefore, Miner's rule, together with a simplified S-N curve, is yet incomplete for the service life estimation.

Several hypotheses were developed to address this issue, e.g. Miner's rule without endurance strength [168, p. 285], Miner's rule modified by Haibach (Miner Haibach hypothesis) [175], Miner's rule modified by Liu and Zenner [176]. They carry over the linear accumulation assumption, but use a modified S-N curve (see [153, pp. 327-329] for a review).

Miner Haibach hypothesis considers the slope of S-N curve on double logarithmic scales in the endurance strength zone to be $2k - 1$. It is adopted for the fatigue accumulation in this work, since it is supported by the experiments, especially in the automotive industry [36, p. 260, 153, p. 328].

8.1.4 Speed4E Powertrain Service Life Estimation

In the rest of this Chapter, "service life", "fatigue" and "S-N curve" are meant for those relate to B_{10} service life.

The bearing and the gearset configuration are schematically illustrated in Figure 8.9. All bearings are labelled in Figure 8.9(a) and their types are listed in Table E. 2 in Appendix E. All gears are labelled in Figure 8.9(b) and their meaning are listed in Table 8.5.

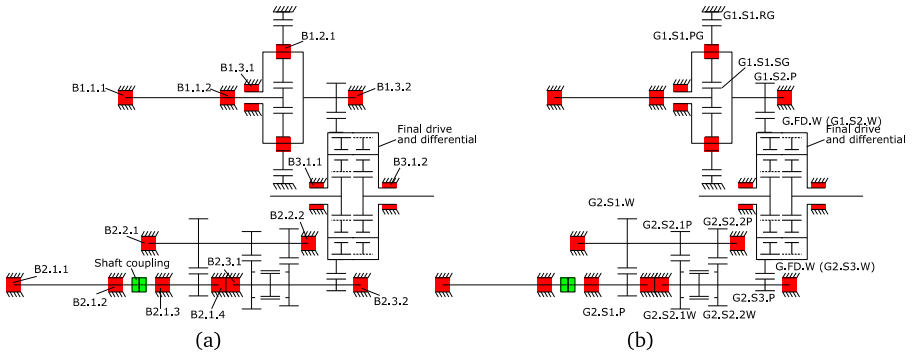


Figure 8.9: Bearing and gearset configuration of Speed4E powertrain.

Table 8.5: The Gears in Speed4E powertrain

| Gear label | Description | Gear label | Description |
|------------|--------------------|------------|---|
| G1.S1.SG | ST1 sun gear | G2.S2.1P | ST2 stage 2 1 st gear pinion |
| G1.S1.PG | ST1 planet gear | G2.S2.1W | ST2 stage 2 1 st wheel |
| G1.S1.RG | ST1 ring gear | G2.S2.2P | ST2 stage 2 2 nd gear pinion |
| G1.S2.P | ST1 stage 2 pinion | G2.S2.2W | ST2 stage 2 2 nd wheel |
| G1.S2.W | ST1 stage 2 wheel | G2.S3. P | ST2 stage 3 pinion |
| G2.S1.P | ST2 stage 1 pinion | G2.S3.W | ST2 stage 3 wheel |
| G2.S1.W | ST2 stage 1 wheel | | |

B1.1.1 and B1.1.2 support the shaft of the EM1 on which G1.S1.SG is mounted. Three planet gears (G1.S1.PG) on three needle bearings (B1.2.1~3) are mounted on the planetary carrier. The ring gear (G1.S1.RG) is fixed on the housing. The planetary carrier and the pinion are supported by B1.3.1 and B1.3.2. B2.1.1 and B2.1.2 support the shaft of the EM2, which is connected to the input shaft of ST2 with a shaft coupling. B2.1.3 and B2.1.4 support the input shaft of the ST2 on which G2.S1.P is mounted. The shaft coupling is considered a torque transferring part. Therefore, B2.1.3 and B2.1.4 are *assumed to withstand all loads on the input shaft of the ST2*. G2.S1.W, G2.S2.1P and G2.S2.2P are mounted on the counter shaft that is supported by B2.2.1 and B2.2.2. G2.S2.1W, G2.S2.2W and G2.S3.P are mounted on the output shaft that is supported by B2.3.1 and B2.3.2. G.FD.W is the wheel that is coupled with G1.S2.P and G2.S3.P that are supported by

B3.1.1 and B3.1.2, respectively. They are part of the final drive and differential. The parameters of the gears and the bearings are listed in Appendix E.

Based on the Miner Haibach hypothesis described in Subsection 8.1.3.3, the modified S-N curves of G1.S1.SG and B1.1.1 are determined and shown in Figure 8.10.

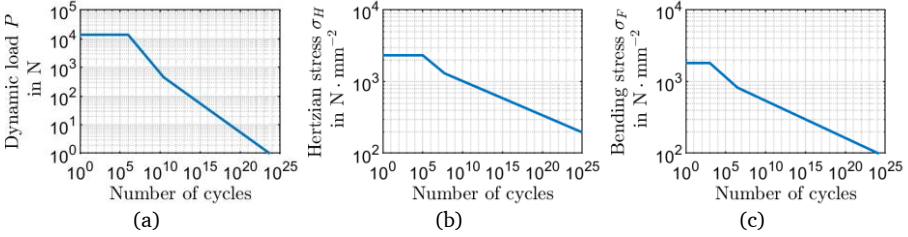


Figure 8.10: Modified S-N Curves of (a) B1.1.1, (b) G1.S1.SG pitting and (c) G1.S1.SG tooth root breakage.

For a gear, σ_F and σ_H are calculated based on (8.2) and (8.3), respectively, in which the normal force F_n and its tangential component F_{tn} are determined by analysing the meshing force that is schematically illustrated in Figure 8.11(a). With the geometry of the gear, namely helix angle α , pressure angle β and pitch diameter, F_n can be decomposed into the transverse force F_t , the radial force F_r and the axial force F_a (see Figure 8.11(b)). F_t is obtained through the torque T_{in} and the pitch diameter, which tells that $F_n \propto T_{in}$ and $F_{tn} \propto T_{in}$. Express F_{tn} as a function of T_{in} and combine all parameters from (8.2) and the geometry of the gear (see Table E. 3) to obtain a coefficient c_F . the relationship between σ_F and T_{in} is

$$\sigma_F = c_F \cdot T_{in}. \quad (8.8)$$

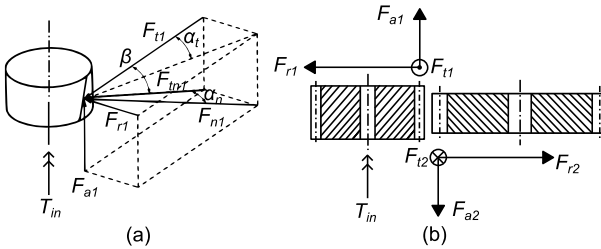


Figure 8.11: Schematic of meshing force

Express F_n as a function of T_{in} and combine all parameters from (8.3) and the geometry of the gear (see Table E. 3) to obtain a coefficient c_H . The relationship between σ_H and T_{in} is

$$\sigma_H = c_H \cdot \sqrt{T_{in}}. \quad (8.9)$$

Forces and torques on bearings need to balance meshing forces of the gears. An example is schematically illustrated in Figure 8.12. The resulting F_y on a bearing is its axial load and the composition of the resulting F_x and F_z is its radial load in (8.4). Express F_a and F_r as two functions of T_{in} and combine all parameters of the gears on the shaft supported by the bearings in question and the spatial locations of all parts (see Table E. 3 and Table E. 4) to obtain a coefficient c_B , the relationship between the equivalent dynamic load L and T_{in} is

$$L = c_B \cdot T_{in}. \quad (8.10)$$

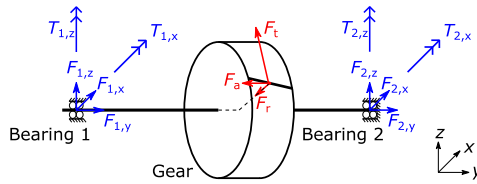


Figure 8.12: An example of forces on bearings

With the service life model presented so far, Figure 8.13 summarizes the accumulated fatigue (blue bars) and the corresponding service lives (orange squares) of the bearings and the gears in Speed4E powertrain on a logarithmic scale. The vehicle drives in the *WLTC* under the control of the predictive EMS. Figure 8.13(a), (b) and (c) present the values of bearing fatigue, tooth root breakage and tooth flank pitting, respectively. In Figure 8.13(b), “pos” and “neg” represent the tooth flank that transmit positive and negative torques, respectively.

In Figure 8.13(a), the accumulated fatigue of B1.1.1 is negligible, since the radial forces on the teeth of G1.S1.SG, the sun gear in the ST1, are balanced and B1.1.2, a fixed bearing, undertakes axial load (see Table E. 2 and Table 8.5 for information of the bearings and the gears). In Figure 8.13(b), the accumulated fatigue of several gears is so small that their service lives exceed the range of the diagram. The values of G2.S2.2P are absent in Figure 8.13(b) and (c), since the ST2 has not engaged in the 2nd gear position in the driving cycle. After comparing Figure 8.13(b) and (c), it can be noticed that the gear service lives w.r.t tooth root breakage are multiple orders of magnitude larger than those w.r.t tooth flank pitting. The simulation result matches the gear design practice mentioned in Section 8.1.3.1 that tooth root breakage is almost eliminated in modern gear design, and it is the tooth flank pitting that is critical to gear service life. *Tooth root breakage is not further discussed in this chapter.* On the other hand, it can be noticed that the service lives of the bearings are multiple orders of magnitude smaller than those of the gears on the same shafts.

The service lives of the bearings limit the powertrain service life of Speed4E powertrain.

B1.1.2 and B2.1.3 have the lowest service life among the bearings in the sub-powertrain 1 (SPT1) and the sub-powertrain 2 (SPT2), respectively. A SPT consists of its corresponding EM and ST. G1.S1.SG and G2.S1.P have the lowest service life among the gears in the SPT1 and the SPT2, respectively. The load spectra of both bearings and the stress spectra of both gears are shown in Figure 8.14 orange bars, while their modified S-N curves are shown as the blue curves.

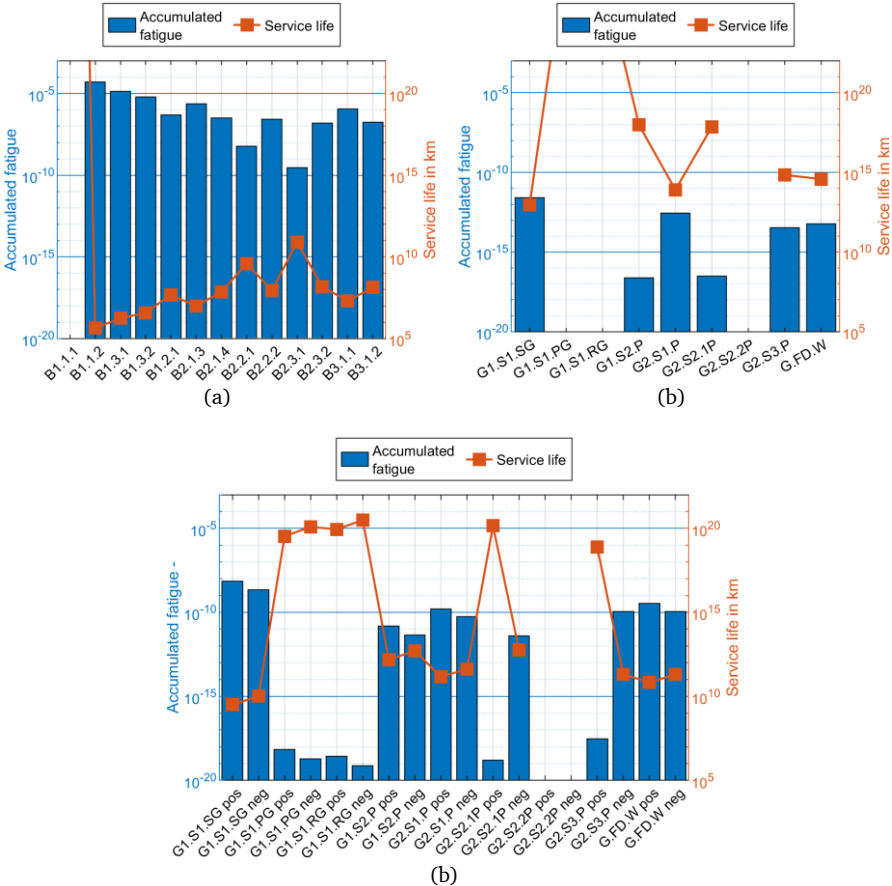


Figure 8.13: Accumulated fatigue after a WLTC and corresponding service lives of the bearings and gears in Speed4E powertrain. (a): Bearing fatigue. (b): Tooth root breakage. (c) Tooth flank pitting.

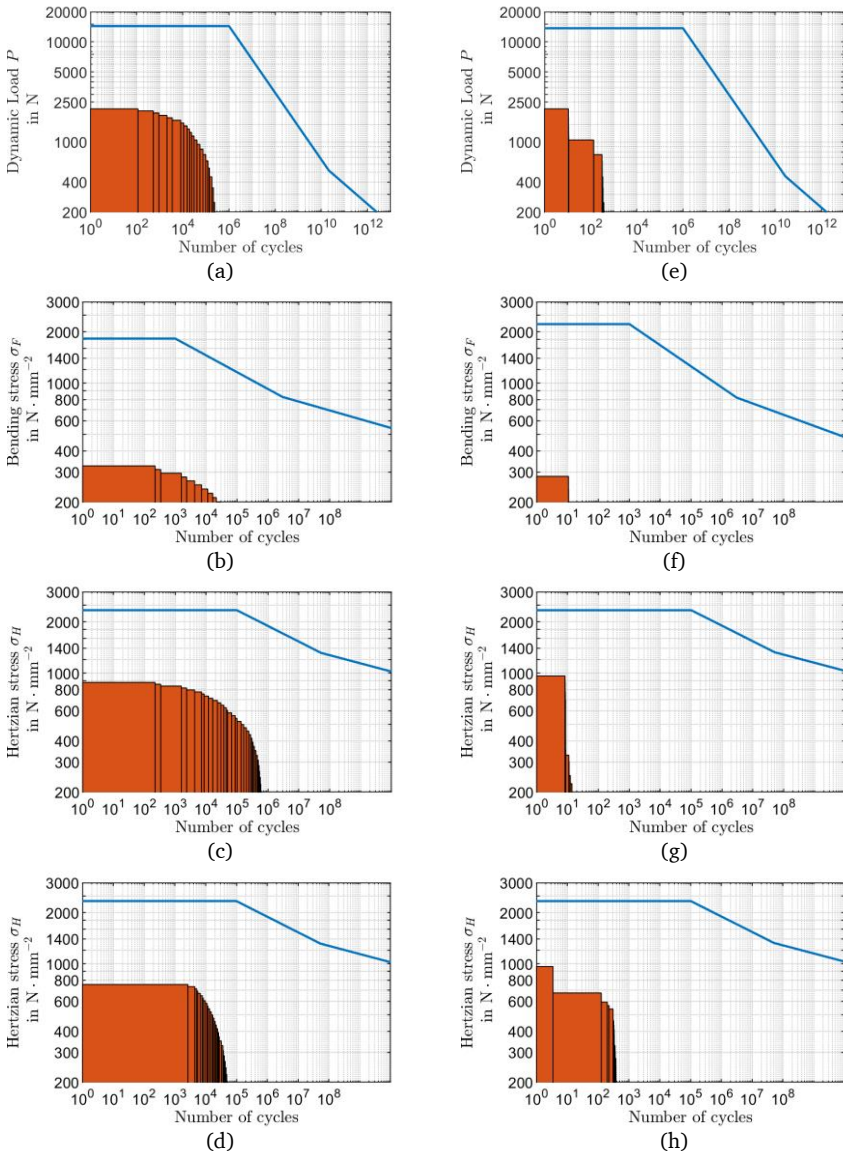


Figure 8.14: Load and stress spectra under the WLTC of (a) B1.1.2, (b) G1.S1.SG breakage, (c) G1.S1.SG pos pitting, (d) G1.S1.SG neg pitting, (e) B2.1.3, (f) G2.S1.P breakage, (g) G2.S1.P pos pitting, and (h) G2.S1.P neg pitting.

It is worth mentioning that the service life of B1.1.2 is the shortest among the bearings of the SPT1, regardless of speed profiles and controls, since they do not change the relationship between the bearing loads and T_1 . For the same reason, the service life of B2.1.3 is the shortest among the bearings of the SPT2. The accumulated fatigue regarding bearings of both SPTs ($D_{B,SPT}$) is

$$D_{B,SPT1} = \min\{D_{B1_x}\} = D_{B1.1.2}, \quad B1_x \in \{B1.1.1, B1.1.2, \dots, B1.3.2\}, \quad (8.11)$$

$$D_{B,SPT2} = \min\{D_{B2_x}\} = D_{B2.1.3}, \quad B2_x \in \{B2.1.3, B2.1.4, \dots, B2.3.2\}. \quad (8.12)$$

For the gears, though the relationship between σ_H and EM torque in (8.3) is nonlinear, they are positively correlated. As a result, the service life of G1.S1.SG remains the shortest among the gears in the SPT1, and the service life of G2.S1.P remains the shortest among the gears in the SPT2. It is affected by gear selection, whether the service life of G2.S2.1P is longer than that of G2.S2.2P. The accumulated fatigue regarding gears of both SPTs ($D_{G,SPT}$) are

$$D_{G,SPT1} = \min\{D_{G1_x}\} = D_{G1.S1.SG}, \quad G1_x \in \{G1.S1.SG, G1.S1.PG, \dots, G1.S2.P\}, \quad (8.13)$$

$$D_{G,SPT2} = \min\{D_{G2_x}\} = D_{G2.S1.P}, \quad G2_x \in \{G2.S1.P, G2.S2.P, \dots, G2.S3.P\}. \quad (8.14)$$

In what follows, only B1.1.2 and B2.1.3 are discussed in terms of bearings, and G1.S1.SG and G2.S1.P in terms of gears. The fatigue of other parts are presented in Appendix H, when necessary.

The service life (distance) SL in km is calculated by

$$SL = \frac{Dist}{D}, \quad (8.15)$$

where $Dist$ denotes the distance that a vehicle travels, while the part operates under the load profile. In the case of the parts listed in Figure 8.13, $Dist$ equals the distance of the WLTC. The service life of the powertrain is limited by the part with the shortest service life, i.e.

$$SL_{PT} = \min\{SL_{G,SPT1}, SL_{G,SPT2}, SL_{B,SPT1}, SL_{B,SPT2}\}. \quad (8.16)$$

8.2 Development of the Multi-Criteria Operation Strategy

As mentioned in Section 8.1.1, the service life of a powertrain depends on the service lives of its parts. The service life of Speed4E powertrain under the *WLTC* equals the one of B1.1.2, the bearing supporting the input shaft of the ST1, since it has the *lowest* service life in the SPT1 and in the powertrain (see Figure 8.13(a)). When its value is reached, B2.1.3, the bearing supporting the input shaft of the ST2 with the *least* service life in the SPT2, has only fatigued for less than 1%. The goal of the MCOS is to *increase the shortest service life in the powertrain, through which the*

powertrain service life is extended. It ought to have the load spectrum of B1.1.2 in Figure 8.14(a) to shrink and that of B2.1.3 in Figure 8.14(c) to expand. As a result, the deviation decreases and the powertrain service life is extended. Meanwhile, the energy consumption can undesirably increase, which causes a trade-off. As shown Section 8.1.4, the bearing service lives limit the Speed4E powertrain service life, which makes the MCOS considering the service lives of the bearings the focus. The one for gears are provided for the applications where the gears are the bottlenecks.

8.2.1 Formulation of the Multi-Criteria HOCP

The multi-criteria HOCP takes over the one in (5.5) with a reformulated cost functional that encodes the goal of the MCOS. The original cost functional is

$$J(q(\cdot), \mathbf{x}(\cdot), \mathbf{u}(\cdot)) = m(\mathbf{x}(t_{f,p})) + \int_{t_{0,p}}^{t_{f,p}} l(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (5.6)$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + PL_{EM1}(\omega_1, u_1) + PL_{EM2}(\omega_2, u_2), \quad (5.7)$$

$$m(\mathbf{x}(t_{f,p})) = \beta \left\| \mathbf{x}(t_{f,p}) - \mathbf{x}_{t_{f,p}} \right\|^2, \quad (5.8)$$

$$\mathbf{x}_{t_{f,p}} = \frac{v_{pred}(t_{f,p})}{r} \cdot (i_1, i_2(q(t_{f,p})))^T, \quad (5.9)$$

which considers solely energy consumption. To align with the minimization nature of the HOCP, the MCOS considers reducing fatigues instead of maximizing service life. J_{multi} , the cost functional of the multi-criteria HOCP, defined as

$$J_{multi} = m(\mathbf{x}(t_{f,p})) + \int_{t_{0,p}}^{t_{f,p}} l(\mathbf{x}(t), \mathbf{u}(t)) dt + \gamma \cdot J_{fatigue}, \quad (8.17)$$

is to be minimized, where $J_{fatigue}$ denotes the fatigue cost functional and γ is the fatigue coefficient.

8.2.1.1 Fatigue Cost Functional for Bearings

$J_{fatigue}$ can be either endpoint or integral functional. Three options are available:

- Option 1: Minimize the deviation between $D_{B,SPT1}$ and $D_{B,SPT2}$, as in

$$J_{fatigue} = \left\| D_{B,SPT1} - D_{B,SPT2} \right\|^2, \quad (8.18)$$

$$D_{B,SPT1} = \int_{t_{0,p}}^{t_{f,p}} d_{B,SPT1}(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (8.19)$$

$$d_{B,SPT1}(\mathbf{x}(t), \mathbf{u}(t)) = \frac{\omega_1}{2\pi \cdot N_{B1.1.2}(P_{B1.1.2}(u_1))} = \frac{\omega_1}{2\pi \cdot N_{B1.1.2}(c_{B1.1.2} \cdot |u_1|)}, \quad (8.20)$$

$$D_{B,SPT2} = \int_{t_{0,p}}^{t_{f,p}} d_{B,SPT2}(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (8.21)$$

$$d_{B,SPT2}(\mathbf{x}(t), \mathbf{u}(t)) = \frac{\omega_2}{2\pi \cdot N_{B2.1.3}(P_{B2.1.3}(u_2))} = \frac{\omega_2}{2\pi \cdot N_{B2.1.3}(c_{B2.1.3} \cdot |u_2|)}, \quad (8.22)$$

where d denotes the fatigue rate function, whose integral is fatigue. $c_{B1.1.2}$ and $c_{B2.1.3}$ are the constant coefficients described in (8.10) for B1.1.2 and B2.1.3. $N_{B1.1.2}$ and $N_{B2.1.3}$ represent the functions that determine the permissible cycle numbers through (8.1). Since the fatigue caused by the positive and negative torques are linearly accumulated, the absolute value of the torques are used. It was reported in [37] that minimizing the deviation of fatigue reduced the largest fatigue until γ reached a certain threshold, beyond which the deviation was further reduced but not the fatigue.

- Option 2: Minimize the average value of $D_{B,SPT1}$ and $D_{B,SPT2}$, as in

$$J_{\text{fatigue}} = \frac{D_{B,SPT1} + D_{B,SPT2}}{2}. \quad (8.23)$$

Due to the exponential correlation of (8.4), the average value depends hugely on the larger accumulated fatigue. Therefore, the larger accumulated fatigue decreases together with the average value, which aligns with the goal.

- Option 3: Minimize the larger one of $D_{B,SPT1}$ and $D_{B,SPT2}$, as in

$$J_{\text{fatigue}} = \max\{D_{B,SPT1}, D_{B,SPT2}\}. \quad (8.24)$$

Option 3 is equivalent to the goal.

Option 2 can be implemented either as an endpoint functional or as an integral functional, while Option 3 needs to be implemented as an endpoint functional, due to the $\max\{\cdot\}$ operation. Though the HOCPs with both kinds of functional are theoretically solvable, a new state vector $\mathbf{x}_{\text{multi}} := (\omega_1, \omega_2, D_{B,SPT1}, D_{B,SPT2})^T$ is mandatory in the case of an endpoint functional. Consequently, the dimension of the new costates is four and a 4-D TPBVP needs to be solved.

The crux of the choice is the convergence of the numerical solutions. Recall that the convergence of the HMP is determined by the transversality condition of the HOCP that is a function of the states at the final time. With Option 2 as an integral functional, there are

$$J_{\text{fatigue}} = \int_{t_{0,p}}^{t_{f,p}} l_{\text{fatigue}}(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (8.25)$$

$$l_{\text{fatigue}}(\mathbf{x}(t), \mathbf{u}(t)) = \frac{d_{\text{B,SPT1}}(\mathbf{x}(t), \mathbf{u}(t)) + d_{\text{B,SPT2}}(\mathbf{x}(t), \mathbf{u}(t))}{2}, \quad (8.26)$$

$$J_{\text{multi}} = m(\mathbf{x}(t_{f,p})) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t)) dt + \gamma \cdot J_{\text{fatigue}} \quad (8.27)$$

$$= m(\mathbf{x}(t_{f,p})) + \int_{t_0}^{t_f} [l(\mathbf{x}(t), \mathbf{u}(t)) + \gamma \cdot l_{\text{fatigue}}] dt. \quad (8.28)$$

The transversality condition is a function of angular velocities, while the time derivatives of the costates are the sum of the original ones in (5.19) and $\gamma \cdot \frac{\partial l_{\text{fatigue}}}{\partial \mathbf{x}_q}$. Large γ can cause extreme difficulty for convergence.

On the other hand, with Option 2 as an endpoint functional and a 4-D states $\mathbf{x}_{\text{multi}}$, there are

$$J_{\text{multi}} = m_{\text{multi}}(\mathbf{x}_{\text{multi}}(t_{f,p})) + \int_{t_{0,p}}^{t_{f,p}} l(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (8.29)$$

$$m_{\text{multi}}(\mathbf{x}(t_f)) = \beta \|\mathbf{x}(t_{f,p}) - \mathbf{x}_{t_{f,p}}\|^2 + \gamma \cdot \frac{D_{\text{B,SPT1}} + D_{\text{B,SPT2}}}{2}. \quad (8.30)$$

Note that $\mathbf{x} := (\omega_1, \omega_2)^T$. The indexed Hamiltonian of the HOCP with J_{multi} is

$$\begin{aligned} \mathcal{H}_{\text{multi},q}(\mathbf{x}_{\text{multi}}, \mathbf{u}, \boldsymbol{\lambda}_{\text{multi},q}) &= \boldsymbol{\lambda}_{\text{multi},q}^T \cdot \dot{\mathbf{x}}_{\text{multi}} + l(\mathbf{x}, \mathbf{u}) \\ &= \lambda_{\text{multi},1,q} \cdot \dot{\omega}_1 + \lambda_{\text{multi},2,q} \cdot \dot{\omega}_2 + \lambda_{\text{multi},3,q} \\ &\quad \cdot d_{\text{B,SPT1}}(\omega_1, u_1) + \lambda_{\text{multi},4,q} \cdot d_{\text{B,SPT2}}(\omega_2, u_2) + l(\mathbf{x}, \mathbf{u}) \\ &= \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q) + \lambda_{\text{multi},3,q} \cdot d_{\text{B,SPT1}}(\omega_1, u_1) \\ &\quad + \lambda_{\text{multi},4,q} \cdot d_{\text{B,SPT2}}(\omega_2, u_2), \end{aligned} \quad (8.31)$$

where $\mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q)$ and $\boldsymbol{\lambda}_q = (\lambda_{1,q}, \lambda_{2,q})^T$ are the indexed Hamiltonian and the costates of the original HOCP described in Section 5.2.2. $\boldsymbol{\lambda}_{\text{multi},q}$ denotes the indexed costates of the HOCP with J_{multi} , whose time derivatives are

$$\begin{aligned} \dot{\boldsymbol{\lambda}}_{\text{multi},q}(t) &= (\dot{\lambda}_{\text{multi},1,q}, \dot{\lambda}_{\text{multi},2,q}, \dot{\lambda}_{\text{multi},3,q}, \dot{\lambda}_{\text{multi},4,q})^T \\ &= -\frac{\partial \mathcal{H}_{\text{multi},q}}{\partial \mathbf{x}_q}(\mathbf{x}_{\text{multi}}^*, \mathbf{u}^*, \boldsymbol{\lambda}_{\text{multi},q}), \end{aligned} \quad (8.32)$$

$$\dot{\lambda}_{\text{multi},1,q} = \dot{\lambda}_{1,q}, \quad (8.33)$$

$$\dot{\lambda}_{\text{multi},2,q} = \dot{\lambda}_{2,q}, \quad (8.34)$$

$$\dot{\lambda}_{\text{multi},3,q} = 0, \quad (8.35)$$

$$\dot{\lambda}_{\text{multi},4,q} = 0. \quad (8.36)$$

$\lambda_{\text{multi},3,q}(\cdot)$ and $\lambda_{\text{multi},4,q}(\cdot)$ are two horizontal lines. Their values can be determined by the transversality condition, which is

$$\lambda_{\text{multi},q}(t_f) = \nabla_{\mathbf{x}_{\text{multi}}} m_{\text{multi}}(\mathbf{x}_{\text{multi}}(t_f)). \quad (8.37)$$

The first two entries of $\lambda_{\text{multi},q}(t_f)$ are the same to $\lambda_q(t_f)$. The third and the fourth entries are

$$\lambda_{\text{multi},3,q}(t_f) = \frac{\gamma}{2}, \quad (8.38)$$

$$\lambda_{\text{multi},4,q}(t_f) = \frac{\gamma}{2}, \quad (8.39)$$

which indicates that

$$\lambda_{\text{multi},3,q}(\cdot) = \frac{\gamma}{2}, \quad (8.40)$$

$$\lambda_{\text{multi},4,q}(\cdot) = \frac{\gamma}{2}. \quad (8.41)$$

Replace $\lambda_{\text{multi},3,q}$ and $\lambda_{\text{multi},4,q}$ in (8.31) to obtain

$$\mathcal{H}_{\text{multi},q}(\mathbf{x}, \mathbf{u}, \lambda_{\text{multi},q}) = \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \lambda_q) + \frac{\gamma}{2} \cdot d_{\text{B,SPT1}}(\omega_1, u_1) + \frac{\gamma}{2} \cdot d_{\text{B,SPT2}}(\omega_1, u_1). \quad (8.42)$$

In summary, the differential equation system with $\mathbf{x}_{\text{multi}}$ and λ_{multi} is equivalent to the one with \mathbf{x} and λ . The transversality condition is the same to the original predictive EMS, which is free from γ . The instantaneous Hamiltonian is changed to $\mathcal{H}_{\text{multi},q}$. Option 2 as an endpoint functional is preferable to that of an integral functional.

The max{·} operation makes the HOCP with Option 3 more difficult to be solved. Only under the assumption that $D_{\text{B,SPT1}}(t_f) = D_{\text{B,SPT2}}(t_f)$ is guaranteed, the HOCP can be shown to be equivalent to Option 2 as an endpoint functional (See Appendix F). The assumption can be interpreted as the MCOS “wishes” to bring $D_{\text{B,SPT1}}$ and $D_{\text{B,SPT2}}$ to the same value.

8.2.1.2 Fatigue Cost Functional for Gears

Replace J_{fatigue} in (8.30) with J'_{fatigue} , the fatigue cost functional for gears,

$$J'_{\text{fatigue}} = \frac{D_{G,\text{SPT1}} + D_{G,\text{SPT2}}}{2} = \frac{\int d_{G,\text{SPT1}} dt + \int d_{G,\text{SPT2}} dt}{2}, \quad (8.43)$$

$$d_{G,\text{SPT1}}(\mathbf{x}(t), \mathbf{u}(t)) = \frac{\omega_1}{2\pi \cdot N_{G1,\text{S1},\text{SG}}(\sigma_{H,G1,\text{S1},\text{SG}}(u_1))} = \frac{\omega_1}{2\pi \cdot N_{G1,\text{S1},\text{SG}}(c_{H,G1,\text{S1},\text{SG}} \cdot \sqrt{u_1})}, \quad (8.44)$$

$$d_{G,\text{SPT2}}(\mathbf{x}(t), \mathbf{u}(t)) = \frac{\omega_2}{2\pi \cdot N_{G2,\text{S1},\text{P}}(\sigma_{H,G2,\text{S1},\text{P}}(u_1))} = \frac{\omega_2}{2\pi \cdot N_{G2,\text{S1},\text{P}}(c_{H,G2,\text{S1},\text{P}} \cdot \sqrt{u_2})}, \quad (8.45)$$

where $c_{H,G1,\text{S1},\text{SG}}$ and $c_{H,G2,\text{S1},\text{P}}$ are the constant coefficients described in (8.8) for G1.S1.SG and G2.S1.P. $N_{G1,\text{S1},\text{SG}}$ and $N_{G2,\text{S1},\text{P}}$ represent the functions that determine the permissible cycle numbers of given Hertzian stress through (8.4) and (8.5).

8.2.2 Solution of the HOCP for the Multi-Criteria Operation Strategy

The HOCP with the fatigue cost functional for bearings is firstly discussed. The HOCP with J_{multi} from (8.29) and (8.30) is solved by the algorithms in Chapter 5. The instantaneous Hamiltonian to be minimized is changed to $\mathcal{H}_{\text{multi},q}$ in (8.42). It is in nature a predictive MCOS. However, since the predictive aspect is less relevant to this chapter, the word “predictive” is omitted.

In the case of the HOCP with the fatigue cost functional for gears, replace $d_{B,\text{SPT1}}$ and $d_{B,\text{SPT2}}$ in (8.42) with $d_{G,\text{SPT1}}$ and $d_{G,\text{SPT2}}$, respectively.

8.3 Effect of Energy Management Strategies on Service Life

Before evaluating the MCOS developed in Section 8.2 in a driving cycle simulation, it is necessary to have a clear view on how a MCOS influence the service lives of the parts. More specifically, how do the torques of both EMs, the gear positions and the shift processes influence the fatigue rate function of different parts?

8.3.1 Torques and Gear Positions

Express ω as a function of the vehicle speed, i.e. $\omega = v \cdot \frac{i}{r}$, and move v to the left side of (8.20), (8.21) (8.22), (8.44) and (8.45) to obtain $\frac{d}{v}$, fatigue rate over vehicle speed. For instance, there is

$$\frac{d_{B,\text{SPT1}}}{v} = \frac{i_1}{2\pi r \cdot N_{B1,1,2}(c_{B1,1,2} \cdot |u_1|)}. \quad (8.46)$$

It evaluates how fast a part fatigues w.r.t. the torques.

As a MCOS changes T_1 and T_2 , $\frac{d}{v}$ of G1.S1.SG, G2.S1.P, B1.1.2 and B2.1.3 change. Figure 8.15

presents the influence of the torques on $\frac{d}{v}$, while the powertrain outputs in total 1000 Nm torque and the ST2 is in the 1st gear position. In each diagram, the x-axis and the y-axis are the torques, and the z-axis is $\frac{d}{v}$ in a logarithmic scale. The z-axes in Figure 8.15(a) and (b) share the same range for a better comparison between the gears and the bearings in question. The z-axis in Figure 8.15(c) is in a smaller range to show a detailed view of $\frac{d}{v}$ of the bearings. T_1 and T_2 are constrained by the inequality constraints, (5.13) and (5.14), as well as the equality constraint, (4.16). The admissible control space of (T_1, T_2) (see Section 5.3.1) forms a black curve in each diagram. Note that $d_G \propto T^{\frac{k}{2}}, T \geq 0, k > 20$ and $d_B \propto |T|^k, k \geq 3$ are convex, in which k is the service exponent of the part in question. The curves in Figure 8.15 seem concave, due to their logarithmic scales.

In Figure 8.15(a), the solid curves and the dashed curves represent the tooth flank “pos” and

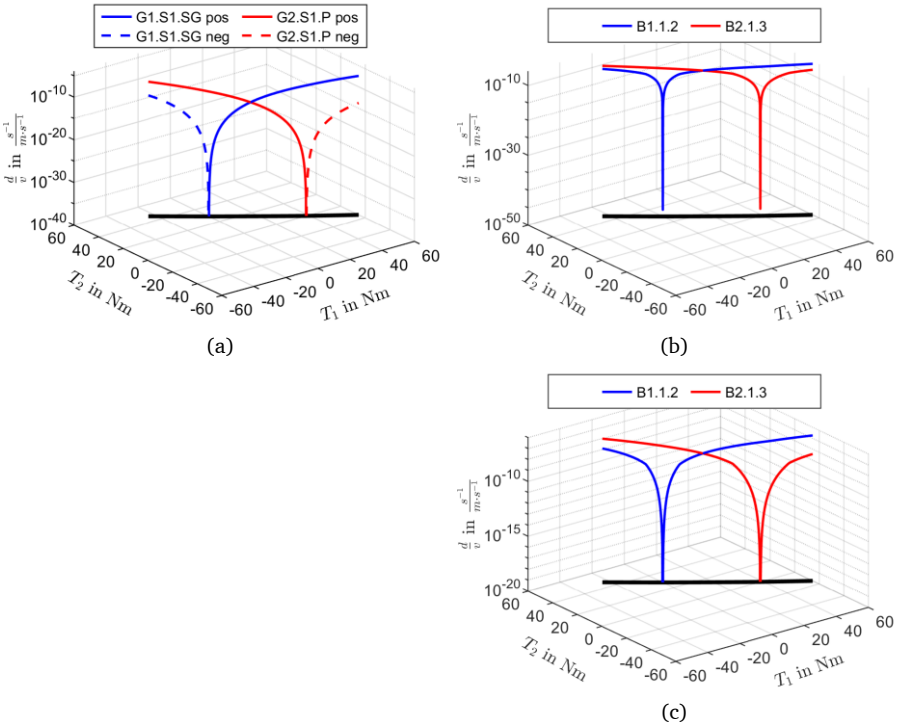


Figure 8.15: Fatigue rate over angular velocity of (a): gears, (b): bearings and (c): bearings zoom in. The total output torque is 1000 Nm. The ST2 in the 1st gear position.

“neg”, respectively. Clearly, either flank fatigues, only when it transmits torques. When total output torque is positive, Lemma 2 and Lemma 3 have shown that the minimal of the power consumption

almost⁶⁰ falls out of the range where $T_1 \cdot T_2 < 0$. Following the same procedure, it can be proved by contradiction that the minimal of the average fatigue rate over vehicle speed, i.e. $0.5 \left(\frac{d_{G,1}(T_1)}{v} + \frac{d_{G,2}(T_2)}{v} \right)$, and the minimal of the highest fatigue rate over vehicle speed, i.e. $\max \left\{ \frac{d_{G,1}(T_1)}{v}, \frac{d_{G,2}(T_2)}{v} \right\}$, fall out of this range⁶¹. The proofs are omitted, since they are repetitions of the proof for Lemma 2. Intuitively, in the range that $T_1 \cdot T_2 < 0$, both functions in question are larger than at the closest point where $T_1 \cdot T_2 = 0$.

As (T_1, T_2) moves from the point that $T_1 = 0$ to the point that $T_2 = 0$ in Figure 8.15(a), $\frac{d_G}{v}$ of G1.S1.SG pos increases and that of G2.S1.P pos decreases. The minimal of the function $\max \left\{ \frac{d_{G,1}(T_1)}{v}, \frac{d_{G,2}(T_2)}{v} \right\}$ is at the point where $\frac{d_{G,1}(T_1)}{v} = \frac{d_{G,2}(T_2)}{v}$, which can be verified by the observation in Figure 8.16, which presents $\max \left\{ \frac{d_{G,1}(T_1)}{v}, \frac{d_{G,2}(T_2)}{v} \right\}$ and $0.5 \left(\frac{d_{G,1}(T_1)}{v} + \frac{d_{G,2}(T_2)}{v} \right)$ in the admissible control space. Furthermore, by observation, the minimal of both functions fall at the same point.

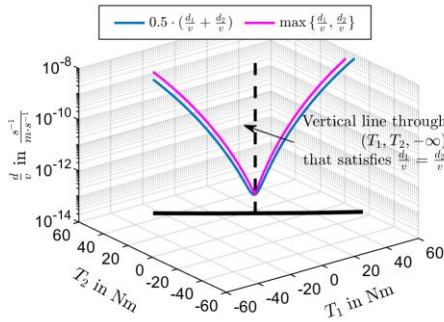


Figure 8.16: The average and the larger values of $\frac{d}{v}$ of G1.S1.SG pos and G2.S1.P pos. The total output torque is 1000 Nm. The ST2 in the 1st gear position.

With a rising γ , the MCOS with the fatigue cost functional for gears shifts the torques towards the point where $\frac{d_1(T_1)}{v} = \frac{d_2(T_2)}{v}$, assuming the discrete control unchanged.

As shown in Figure 8.15(b), $\frac{d_B}{v}$ of B1.1.2 and B2.1.3 are significantly larger than those of the

⁶⁰ When the total output torque in a small range, it can happen that $T_1 \cdot T_2 < 0$. See Section 5.3.2.

⁶¹ Prove by contradiction: Suppose for the purpose of contradiction that a point (T_1, T_2) with $T_1 \cdot T_2 < 0$ minimizes $\max \left\{ \frac{d_1}{v}, \frac{d_2}{v} \right\}$. Show that, with a small increment of the negative torque towards zero, the new point finds a smaller $\max \left\{ \frac{d_1}{v}, \frac{d_2}{v} \right\}$. Same can be done for $0.5 \left(\frac{d_1}{v} + \frac{d_2}{v} \right)$.

gears just mentioned. The change rate of the logarithmic value of $\frac{d_{B,1}(T_1)}{v}$ jumps at $T_1 \approx -8$ Nm and $T_1 \approx 8$ Nm in Figure 8.15(c). The dynamic loads of B1.1.2 corresponding these torque values are the fatigue limit loads. The same situation happens to B2.1.3, whose fatigue limit loads correspond to $T_2 \approx \pm 9$ Nm. It can be verified by Figure 8.14(a) and (c). On the other hand, when transmitting the admissible torque, all gears in Speed4E powertrain operate in their endurance strength zone. Therefore, the derivative of the logarithmic value of $\frac{d_{G,1}(T_1)}{v}$ and $\frac{d_{G,2}(T_2)}{v}$ are continuous.

The discussion w.r.t. the minimal of the average and the larger value of $\frac{d}{v}$ can be carried over. A similar result is obtained: The minimal of both functions of B1.1.2 and B2.1.3 falls at a point (T_1, T_2) that satisfies $\frac{d_{B,1}(T_1)}{v} = \frac{d_{B,2}(T_2)}{v}$, which can be verified by the observation in Figure 8.17, a counterpart of Figure 8.16. Please be aware that $\frac{d_{B,1}(T_1)}{v} = \frac{d_{B,2}(T_2)}{v}$ and $\frac{d_{G,1}(T_1)}{v} = \frac{d_{G,2}(T_2)}{v}$ are not satisfied simultaneously. With a rising γ , the MCOS with the fatigue cost functional for bearings shifts the torques towards the point where $\frac{d_1(T_1)}{v} = \frac{d_2(T_2)}{v}$, assuming the discrete control unchanged.

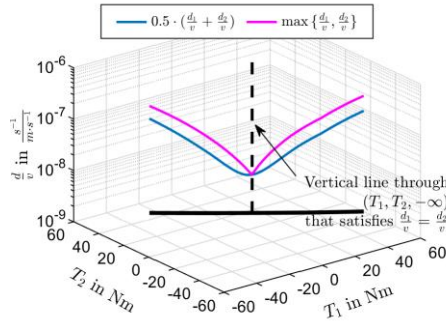


Figure 8.17: The average and the larger values of $\frac{d}{v}$ of B1.1.2 and B2.1.3. The total output torque is 1000 Nm. The ST2 in the 1st gear position.

Figure 8.18 presents the influence of the torques on $\frac{d}{v}$. The total output torque is 1000 Nm and the ST2 is in the 2nd gear position. All settings are the same as in Figure 8.15. Since i_2 changes to the value in the 2nd gear position, the admissible control space (the black curve) changes. Under the equality constraint (4.16), fix T_1 and the output torque, the EM2 needs to provide more torque when $G = 2$ than $G = 1$. Consequently, $\frac{d_{G,2}(T_2)}{v}$ increases. Recall that the minimal of $\max\left\{\frac{d_{G,1}(T_1)}{v}, \frac{d_{G,2}(T_2)}{v}\right\}$ and $0.5\left(\frac{d_{G,1}(T_1)}{v} + \frac{d_{G,2}(T_2)}{v}\right)$ are at the point where $\frac{d_{G,1}(T_1)}{v} = \frac{d_{G,2}(T_2)}{v}$. The increase of $\frac{d_{G,2}(T_2)}{v}$ cause both minimal to rise, which can be verified by the observation of Figure 8.19(a) and (b), the counterparts of Figure 8.16 and Figure 8.17, respectively.

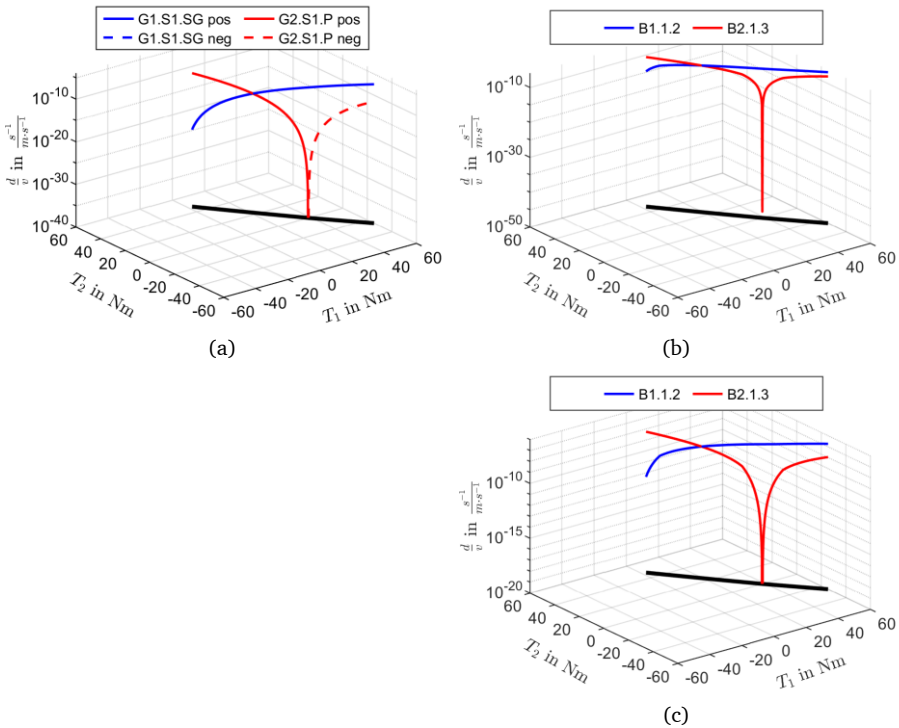


Figure 8.18: Fatigue rate over angular velocity of (a): gears, (b): bearings and (c): bearings zoom in. The total output torque is 1000 Nm. The ST2 in the 2nd gear position.

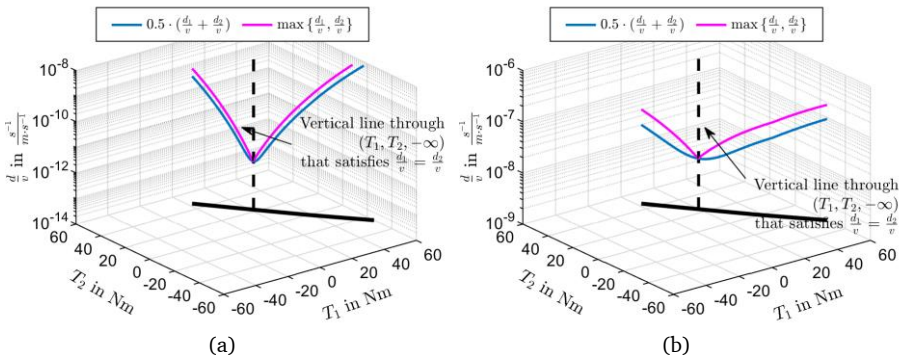


Figure 8.19: The average and the larger values of $\frac{d}{v}$. (a): G1.S1.SG pos and G2.S1.P pos. (b): B1.1.2 and B2.1.3.

8.3.2 Shift Processes

Two different operation strategies generate two different discrete state sequences for a simple speed profile in Figure 8.20. To show the influence of shift processes on service life, the accumulated fatigue between $t = 12$ s and $t = 18$ s is compared against each other.

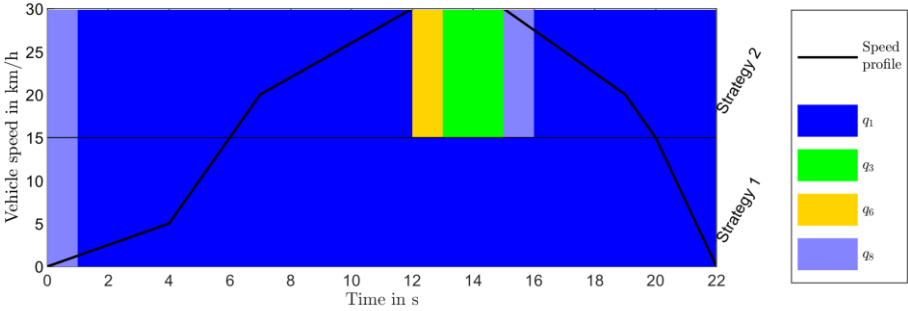


Figure 8.20: Two discrete state sequences.

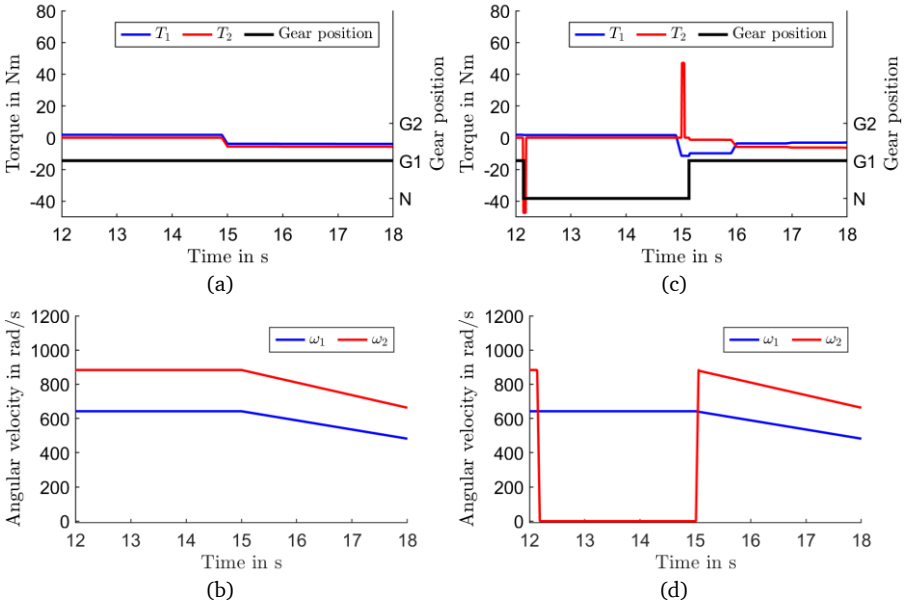


Figure 8.21: The continuous controls, the gear position and the continuous states between 12 s and 18 s. (a) and (b): the strategy 1. (c) and (d): the strategy 2.

Figure 8.21 compares the torques, the gear positions and the EM angular velocities of both strategies between $t = 12$ s and $t = 18$ s. After 3 seconds of constant speed, the vehicle starts to decelerate at $t = 15$ s. Figure 8.21(a) and (b) show the operation of both EMs with the ST2 stays in its 1st gear position. Figure 8.21(c) shows that, with the strategy 2, the EM2 generates its minimal torque and its maximal torque during the shift processes at $t = 12$ s and $t = 15$ s, respectively, to change ω_2 to its target values in the target gear positions in the respective shift processes (see Section 4.4.2 for the shift processes). Figure 8.21(d) shows the corresponding angular velocities.

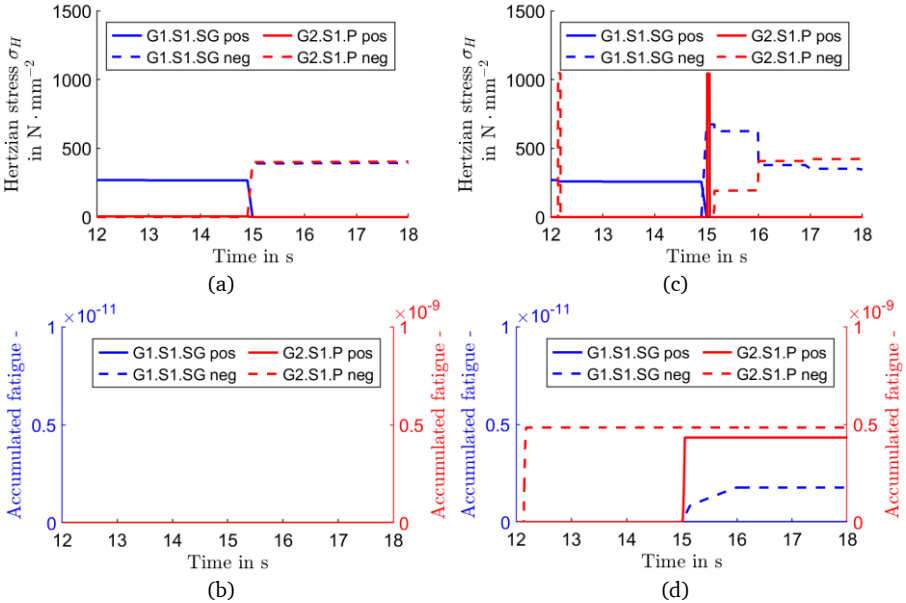


Figure 8.22: The Hertzian stresses on the gears and the resulting fatigue between 12 s and 18 s. (a) and (b): the strategy 1. (c) and (d): the strategy 2.

The Hertzian stresses on G1.S1.SG and G2.S1.P—the solid curves for the tooth flanks “pos”, and the dashed curves for the tooth flanks “neg”—as well as the resulting accumulated fatigue between $t = 12$ s and $t = 18$ s are presented in Figure 8.22. All y-axes are in logarithmic scale. Note that the accumulated fatigue of the gears in Figure 8.22(b) and (d) is associated with different axes according to their colours. The curves of Hertzian stresses in Figure 8.22(a) and (c) reflect the torques in Figure 8.21(a) and (c). Most noticeably, the Hertzian stress spikes on G2.S1.P pos and G2.S1.P neg in Figure 8.22(c) caused far severer fatigue compared to the milder Hertzian stresses: The accumulated fatigue in Figure 8.22(b) is hardly noticeable; The accumulated fatigue of G2.S1.P pos hardly changes after the jump at $t = 15$ s. It is caused by the negative exponent with a large magnitude in (8.1) for the permissible cycle numbers of gears and, consequently, a large

exponent is implicit in (8.44) for the fatigue rate. For the same reason, G1.S1.SG neg fatigues between 15 s and 16 s more than between 16 s and 18 s in Figure 8.22(d). However, it fatigues multiple orders of magnitude less than G2.S1.P.

The dynamic loads on B1.1.2 and B2.1.3 as well as the resulting accumulated fatigue between $t = 12$ s and $t = 18$ s are presented in Figure 8.23. All y-axes are in logarithmic scale. Comparing Figure 8.23(c) and (d), the load spikes cause jumps in accumulated fatigue. However, their ratios to the accumulated fatigue caused by the milder loads, e.g. those between 15 s and 18 s, are much less than the ratios of the accumulated fatigue caused by the stress spikes to that caused by the milder stresses in the case of the gears.

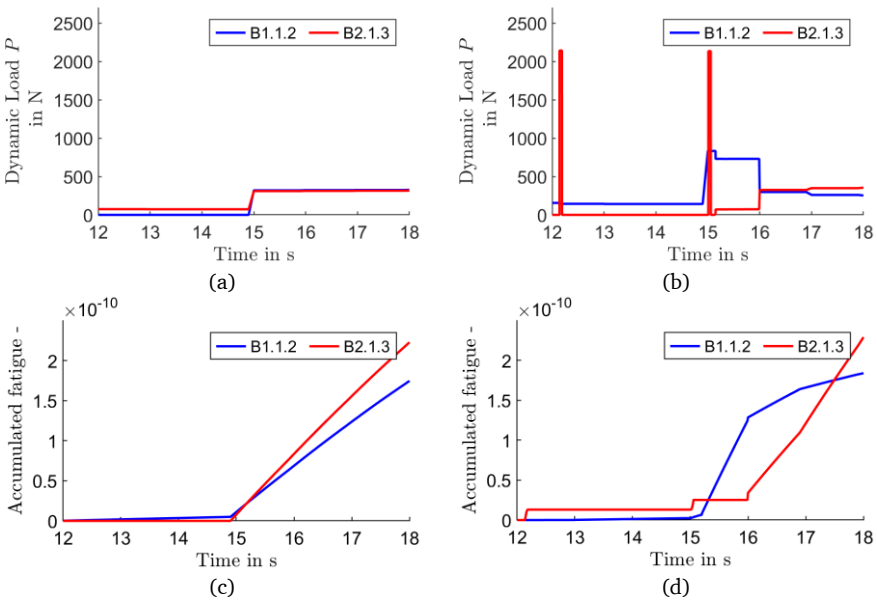


Figure 8.23: The dynamic loads on the bearings and the resulting fatigue between 12 s and 18 s. (a) and (b): the strategy 1. (c) and (d): the strategy 2.

8.4 Driving Cycle Simulation

The MCOS is to be evaluated in the WLTC driving cycle simulation as described in Section 6.3. More specifically, the results of a set of driving cycle simulations of the MCOS with different γ in (8.29) are examined, since it indicates how much the MCOS “wishes” to reduce the average accumulated fatigue so that the service life of the powertrain is extended. The predictive horizon of the MCOS is set to 12, based on the results of the predictive EMS with different predictive horizon in

Section 6.3.5.

It is noteworthy that the multi-criteria HOCF with the speed profile as *a priori* can also be solved by the DP to obtain a global optimal solution. Such a reference MCOS has been investigated in the previous work of the author [37]. Thanks to the known speed profile, it outperformed a MCOS in the form of Option 1 in Section 8.2.1.1. Similar comparison is omitted in this chapter.

For compactness, “energy consumption”, “accumulated fatigue” and “service life” in the following subsections refer to their values subject to the WLTC without explicitly stating it.

The results from Figure 8.13 and Section 8.3.1 have shown that the service life of Speed4E powertrain is limited by those of the bearings, i.e. $SL_{PT} = \min\{SL_{B,SPT1}, SL_{B,SPT2}\}$. The following subsections focus on the MCOS with the fatigue cost functional for bearings. The results of the counterpart for gears from Section 8.2.1.2 are presented in Appendix H. The main difference is that it does not manage to extend the powertrain service life as far as the one for the bearings, since the gear service life is not the bottleneck for the powertrain service life.

8.4.1 Overall Evaluation

Figure 8.24 presents the values of the sampled γ for the evaluation on a logarithmic scale with the index of γ as the x-axis (See Appendix G for the exact values). The indexed γ is used as the fatigue coefficient in the MCOS with the same index. It can be noticed that less amount of γ are sampled, when $\gamma < 10^{10}$ or $\gamma \geq 10^{12}$. In either case, $\gamma \cdot J_{\text{fatigue}}$ is either so much larger or so much smaller than $\int ldt$ that the minimization of (8.17) is insensitive to γ .

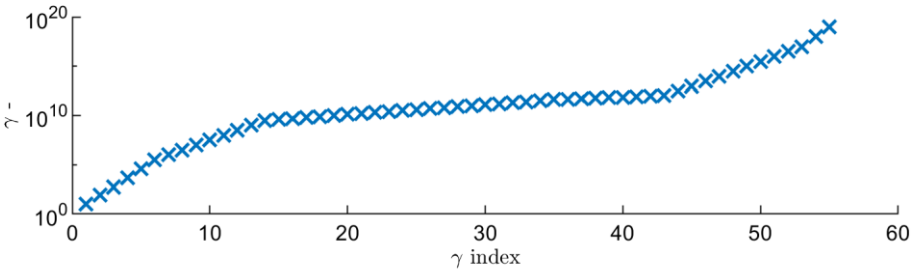


Figure 8.24: The values of γ sampled for the simulation study

Figure 8.25 summarizes the vehicle energy consumption and the average accumulated fatigue of a set of WLTC driving cycle simulations. Note that the average accumulated fatigue is the fatigue cost functional. Each circle represents an indexed simulation with an indexed MCOS. For instance, the circle under the annotation “ γ_{33} ” shows that, under the control of the MCOS with γ_{33} , the

vehicle consumes energy at a rate of 12.46 kWh/100 km, while the average value of the accumulated fatigues of the SPT1 and the SPT2 regarding bearings is ca. 10^{-5} .

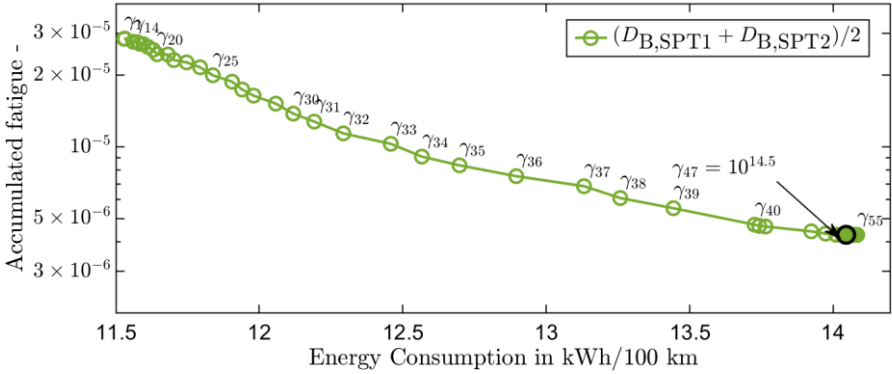


Figure 8.25: The simulation results of the MCOS with the indexed γ : Average accumulated fatigue vs. Energy consumption.

As γ increases in Figure 8.25, the vehicle consumes more energy, while the average accumulated fatigue declines. Due to the sensitivity of the HOCP to γ that has been discussed earlier, plenty of the data points assemble in the range of $\gamma \in [\gamma_1, \gamma_{20}] \cup [\gamma_{40}, \gamma_{50}]$, notwithstanding the larger exponentially increment of γ . In the case of $\gamma \geq \gamma_{47}$, the average accumulated fatigue saturates at 4.23×10^{-6} . Between γ_{49} and γ_{55} , both the energy consumption and the average accumulated fatigue remain unchanged. They are equivalent to the result of a predictive operation strategy considering solely service life.

Figure 8.26 shows the powertrain service life, i.e. $\min\{SL_{B,SPT1}, SL_{B,SPT2}\}$, of the same set of simulations on a logarithmic scale. Through increasing γ and, consequently, minimizing the average accumulated fatigue, the powertrain service life extends. The design of the multi-criteria functional realizes the goal of the MCOS. The MCOS with γ_{49} in Figure 8.26 is equivalent to a predictive operation strategy considering solely service life. Compared to the predictive EMS (equivalent to the MCOS with γ_1), it consumes 22.12 % more energy (from 11.52 kWh/100 km to 14.08 kWh/100 km) but extends the powertrain service life by more than 9 times (from 4.1×10^5 km to 4.24×10^6 km). It is noticed that $\log \frac{SL_{PT}}{1 \text{ km}}$ increases almost linearly with a slope of $76.75 \frac{1}{\text{kWh}/100 \text{ km}}$ before γ_{31} , while the slope changes to $1.25 \frac{1}{\text{kWh}/100 \text{ km}}$ between γ_{31} and γ_{50} . At the elbow point γ_{31} , the vehicle consumes 5.72 % more energy (from 11.52 kWh/100 km to 12.19 kWh/100 km) but extends the powertrain service life by more than three times (from

⁶² The argument of a transcendental function, e.g. a logarithmic function, should be dimensionless.

4.10×10^5 km to 1.80×10^6 km), compared to the predictive EMS.

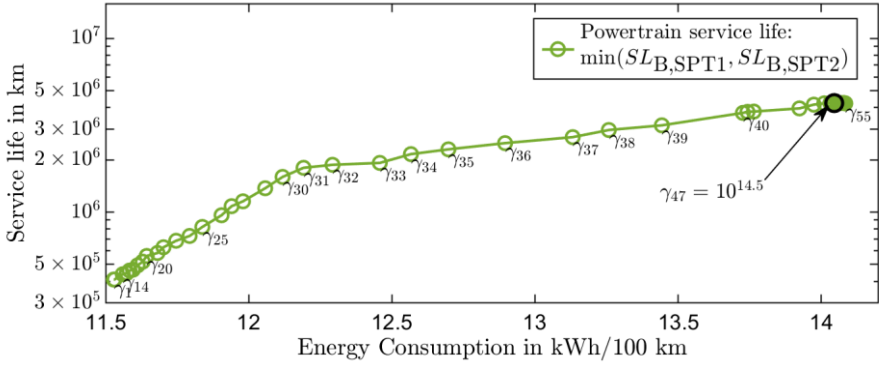


Figure 8.26: The simulation results of the MCOS with the indexed γ : Powertrain service life vs. Energy consumption.

8.4.2 A Closer Look at the Bearings

Figure 8.27 presents the accumulated fatigue of B1.1.2 (the blue curve with circles) and B2.1.3 (the red curve with circles) on a logarithmic scale. They determine $D_{B,SPT1}$ and $D_{B,SPT2}$, as discussed in Section 8.1.4. In the range of $\gamma \in [\gamma_1, \gamma_{12}]$, the data points of B1.1.2 nearly coincide, due to the unsensitivity discussed earlier. So do the data points of B2.1.3. When $\gamma \geq \gamma_{13}$, $\log \frac{D_{B1.1.2}}{1 \text{ km}}$ decreases gradually. On the other hand, $D_{B2.1.3}$ doubles, and $\log \frac{D_{B2.1.3}}{1 \text{ km}}$ increases by 1, as γ rises to γ_{14} , while the energy consumption changes little to none. This will be discussed together with the gears in Section 8.4.3. As γ rises further, $\log \frac{D_{B2.1.3}}{1 \text{ km}}$ changes with a derivative w.r.t. energy consumption that

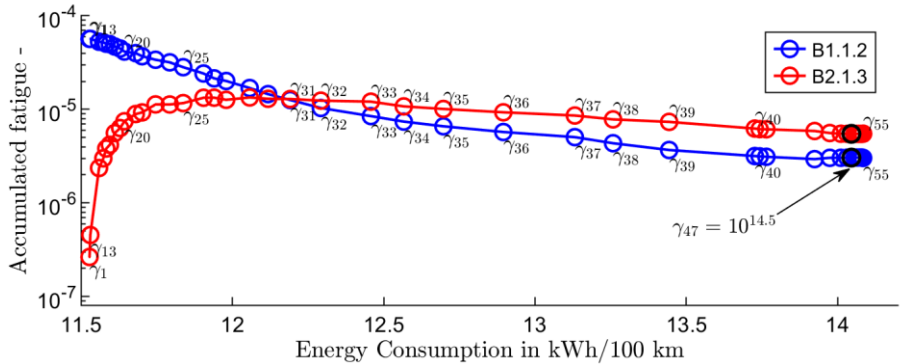


Figure 8.27: The simulation results of the MCOS with the indexed γ : Bearing accumulated fatigue vs. Energy consumption

decreases gradually from a positive to a negative value. $D_{B2.1.3}$ reaches its maximum at γ_{29} . With $\gamma \geq \gamma_{47}$, $D_{B2.1.3}$ remains almost unchanged, while the energy consumption increases slowly.

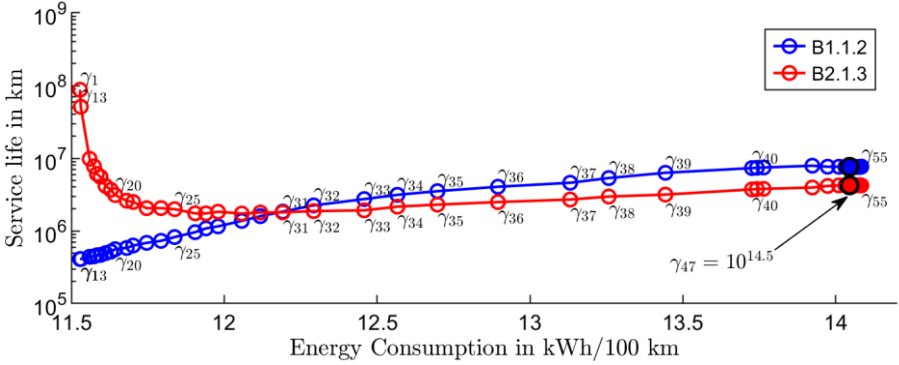


Figure 8.28: The simulation results of the MCOS with the indexed γ : Bearing service life vs. Energy consumption.

Figure 8.28 presents the service lives of B1.1.2 (the red curve with circles) and B2.1.3 (the blue curve with circles). They determine $SL_{B,SPT1}$ and $SL_{B,SPT2}$, as discussed in Section 8.1.4. The image of the curves in Figure 8.28 are the mirror images of those in Figure 8.27 over a horizontal line, since $SL \propto \frac{1}{D}$ by (8.15) and both y-axes are in logarithmic scale. The curve of the powertrain service life in Figure 8.26 is the combination of the curve $SL_{B1.1.2}$ when $\gamma \in [\gamma_1, \gamma_{31}]$ and $SL_{B2.1.3}$ when $\gamma \in (\gamma_{31}, \gamma_{50}]$, which causes two different slopes in Figure 8.26.

Figure 8.29 summarizes the gear distribution under the MCOS with γ_1 , γ_{31} and γ_{55} . Figure 8.30 summarizes the operating duration of both EMs.

As γ changes from γ_1 to γ_{31} , the ST2 operates in the 1st gear position in a larger area compared to the neutral gear position (see Figure 8.29(a) and (b)), since the MCOS has the SPT2 take over part of the loads. Compare Figure 8.30(b) to Figure 8.30(a): With γ_{31} , the EM1 operates only in the range of $T_1 \in [-20, 20]$. On the other hand, the EM2 provides the necessary torques, as T_1 reduces (see Figure 8.30(e)). As a result, most noticeably, the loads with a dynamic load of 2100 N do not appear in the load spectrum of B1.1.2 in Figure 8.31(b), while the load spectrum of B2.1.3 expands at different levels of dynamic load, e.g. 2500 N and 2100 N.

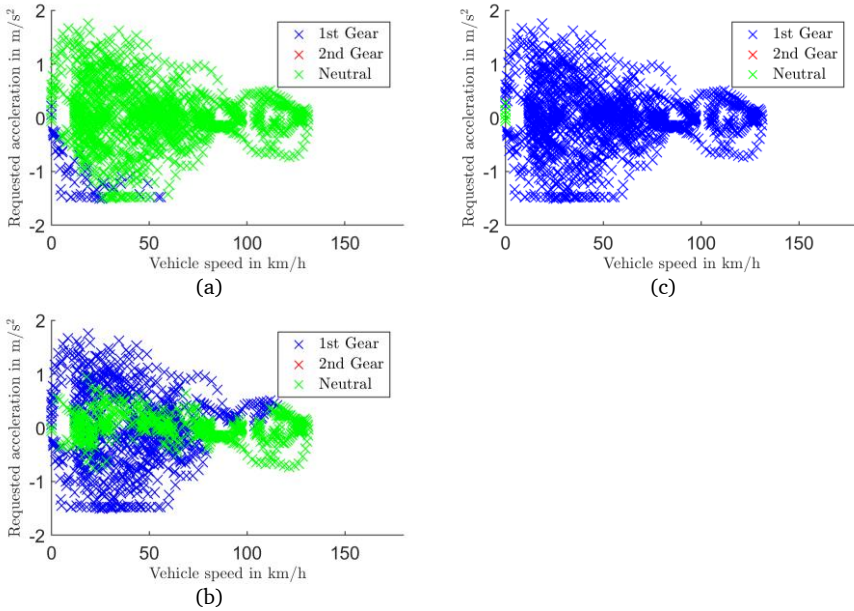


Figure 8.29: Gear distribution of the predictive MCOS with (a) γ_1 , (b) γ_{31} and (c) γ_{55} , respectively.

Figure 8.29(c) shows that the ST2 stays only in the 1st gear position, unless the vehicle speed is close to zero and the acceleration is small. Through avoiding shifts, with γ_{55} , the MCOS has the EM2 hardly operate with its maximum or minimum torques (see Figure 8.30(f)). Furthermore, the MCOS has both SPTs divide the loads as even as possible. Compare Figure 8.30(c) to Figure 8.30(b): With γ_{55} , the EM1 operates only in the range of $T_1 \in [-10, 10]$; The operating duration at the points with small torques rises. On the other hand, the area of operating points of the EM2 narrows in the direction of torque and elongates in the direction of rotational speed (see Figure 8.30(f)), and the operating duration at the area with small torques rises.

The load spectra of B1.1.2 and B2.1.3 change accordingly (see Figure 8.31). The change of the load spectra from γ_1 to γ_{31} clearly follows the change of the operating duration of both EMs discussed earlier. To understand the change of the load spectra from γ_{31} to γ_{55} , the load spectra with γ_{37} are provided additionally in Figure 8.31(c) and (g). As γ increases, the MCOS requires both EMs divide the total output torques more evenly so that the average accumulated fatigue can be further reduced. As a result, there are less instances that either EM provides a major part of the output torque and the other EM complements it. The bars of large dynamic loads in Figure 8.31(c) and (g), therefore, narrow, in compared to Figure 8.31(b) and (f). For instance, those of 1500 N

and 1300 N of B1.1.2 as well as those of 2500 N and 2100 N of B2.1.3. At the same time, the rest of the bars widen. These changes further develop until the load spectra fall in the state as shown in Figure 8.31(d) and (h). Another reason is related to the shift processes, which will be discussed together with the gears in Section 8.4.4.

See Appendix H for the accumulated fatigue and the service lives of all bearings and gears at γ_{31} and γ_{55} .

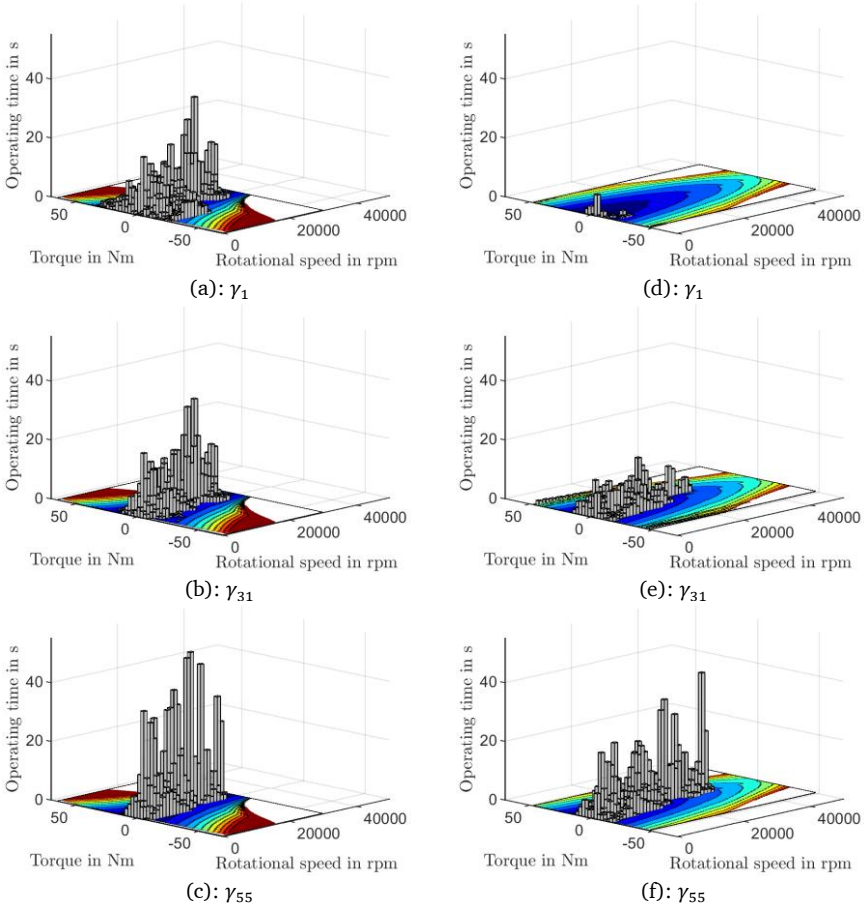


Figure 8.30: Operating duration of both EMs with different γ : (a), (b) and (c): the EM1. (d), (e) and (f): the EM2.

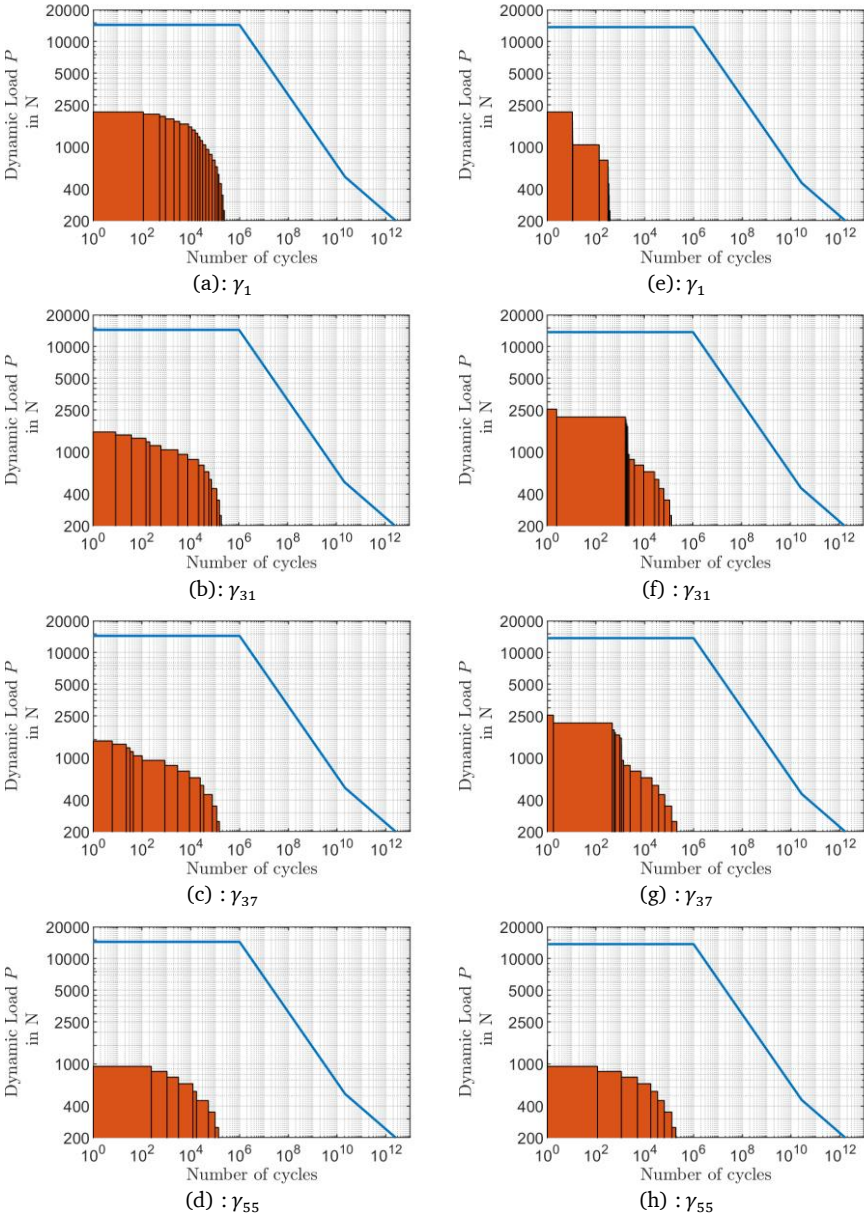


Figure 8.31: Load spectra of B1.1.2 (left side) and B2.1.3 (right side) with different γ .

8.4.3 A Closer Look at the Gears

As γ increases, the gears fatigue differently. Figure 8.32 presents the service life of G1.S1.SG, i.e. $SL_{G1.S1.SG}$ (the blue curves), and that of G2.S2.P, i.e. $SL_{G2.S1.P}$ (the red curves), on a logarithmic scale, where the circles represent the tooth flanks “pos” and the diamonds represent the tooth flanks “neg”. $SL_{G2.S1.P}$ drops to the level of $SL_{G1.S1.SG}$ at γ_{18} , while the similar situation happens with the bearings at γ_{31} , for the fatigue of the gears is more sensitive to the torques, which has been discussed in Section 8.3. As γ further increases, $SL_{G2.S1.P}$ decreases until γ_{29} and rises thereafter. The service lives of both gears saturate after γ reaches γ_{46} . Between γ_{46} and γ_{55} , the energy consumption increases slowly. The relationship between the service life curves and the accumulated fatigue curves has been discussed in 8.4.2. See Figure H. 3 in Appendix H for the accumulated fatigue of the gears.

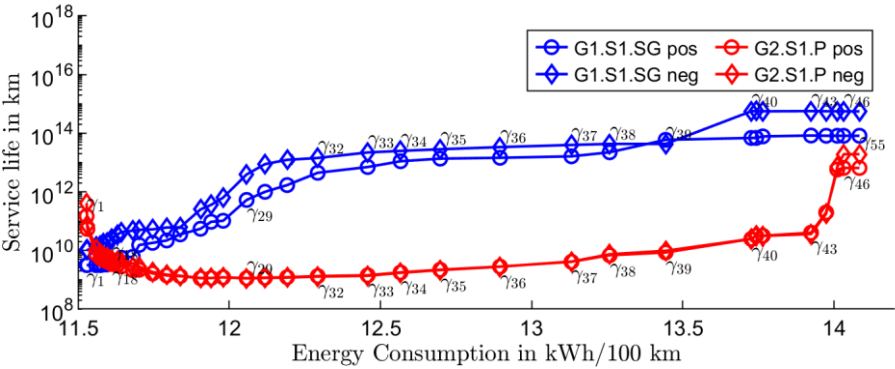


Figure 8.32: The simulation results of the MCOS with the indexed γ : Gear service life vs. Energy consumption.

Most noticeably, after γ_{43} , $SL_{G2.S1.P}$ increases drastically until saturation at γ_{46} . This is mainly caused by the change of the shift count from 42 at γ_{43} to 5 at γ_{45} . The shift counts out of all driving cycle simulations are summarized in Figure 8.33. The positive correlation between the accumulated fatigue of G2.S1.P and the shift count in the driving cycle is presented in Figure 8.34, where the y-axis is in a logarithmic scale. The data points in terms of $D_{G2.S1.P \text{ pos}}$ are omitted, since they mostly coincide with $D_{G2.S1.P \text{ neg}}$ until γ_{45} in Figure 8.32.

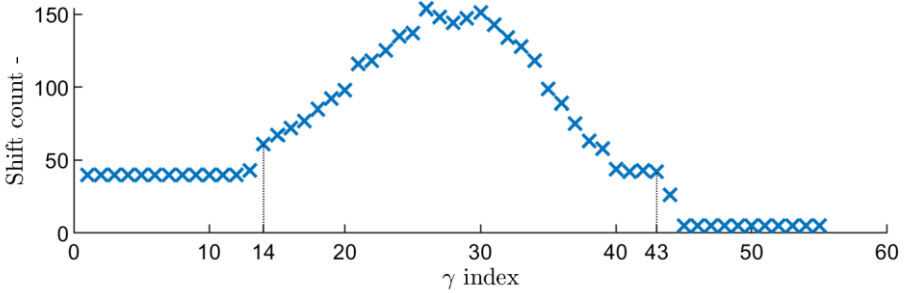


Figure 8.33: Shift counts of all driving cycle simulations.

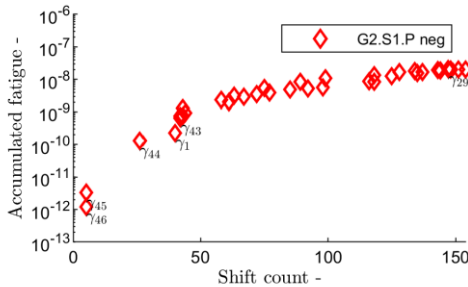


Figure 8.34: The simulation results of the MCOS with the indexed γ : Gear accumulated fatigue vs. Shift count.

The drastic change of $SL_{G2.S1.P}$ from γ_{43} to γ_{45} is caused not only by the shift count but also by the vehicle speed when the shifts happen. Comparing the gear distribution with γ_{43} to that with γ_{45} (Figure 8.35(a) and (b)), it can be noticed that the shift processes with γ_{45} only take place with vehicle speed either close to zero or slightly larger than zero, while those with γ_{43} take place with larger vehicle speed. With lower vehicle speed, the EM2 accelerates and decelerates for extremely short period with its maximal and minimal torque, and the EM2 angular velocities during these processes are low. The fatigue rates in these cases integrate to low values. It is verified by the stress spectra of $SL_{G2.S1.P \text{ neg}}$ with γ_{43} and γ_{45} in Figure 8.36—Note that Figure 8.36(b) is slightly shifted to show the bar with $1050 \text{ N}\cdot\text{mm}^{-2}$.

Figure 8.34 shows that, as γ changes from γ_{45} to γ_{46} , the shift count does not change, but $D_{G2.S1.P \text{ neg}}$ does. Comparing the gear distribution with γ_{45} to that with γ_{46} (Figure 8.35(b) and (c)), it can be noticed that the shift processes from the 2nd gear position to the neutral gear position in Figure 8.35(b) cease to happen. Consequently, $G2.S2.P \text{ neg}$ is free from the stresses caused by

the negative torque spikes at the beginning of these shift processes. It is verified by the disappearance of the narrow bar with $962 \text{ N}\cdot\text{mm}^{-2}$ in Figure 8.36(c). However, the MCOS with γ_{46} still requests the shift processes from the neutral gear position to the 1st gear position. $SL_{G2.S1.P \text{ pos}}$ stays at the same level, as γ changes from γ_{45} to γ_{46} .

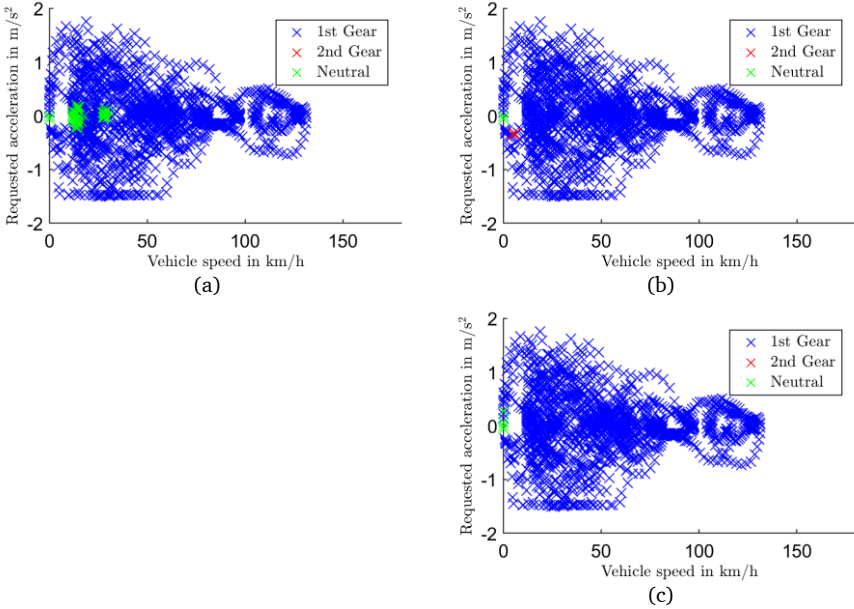


Figure 8.35: Gear distribution of the MCOS with (a) γ_{43} , (b) γ_{45} and (c) γ_{46} .

On the other end of the γ index, as the shift count rises from 43 at γ_{13} to 61 at γ_{14} , $SL_{B2.1.3}$ in Figure 8.28 and $SL_{G2.S1.P \text{ neg}}$ in Figure 8.32 halve. The change of $SL_{B2.1.3}$ is more obvious, for Figure 8.27 is in a smaller scale. They are caused mainly by the torque spikes during the shift processes, which can be verified by the widening of the bars with $962 \text{ N}\cdot\text{mm}^{-2}$ in Figure 8.36(d) and (e), as well as those with 2100 N in Figure 8.37(a) and (b).

As γ rises from γ_{38} to γ_{40} , $SL_{G1.S1.SG \text{ neg}}$ rises above $SL_{G1.S1.SG \text{ pos}}$, since the operating duration of the EM1 with negative torque reduces, which can be observed by comparing Figure 8.38(a) and (b). Due to the logarithmic scale in Figure 8.32, the change of $SL_{G2.S1.P \text{ neg}}$ is unnoticeable, though the EM2 provides the negative torque that compensates for the reduced negative torque of the EM1.

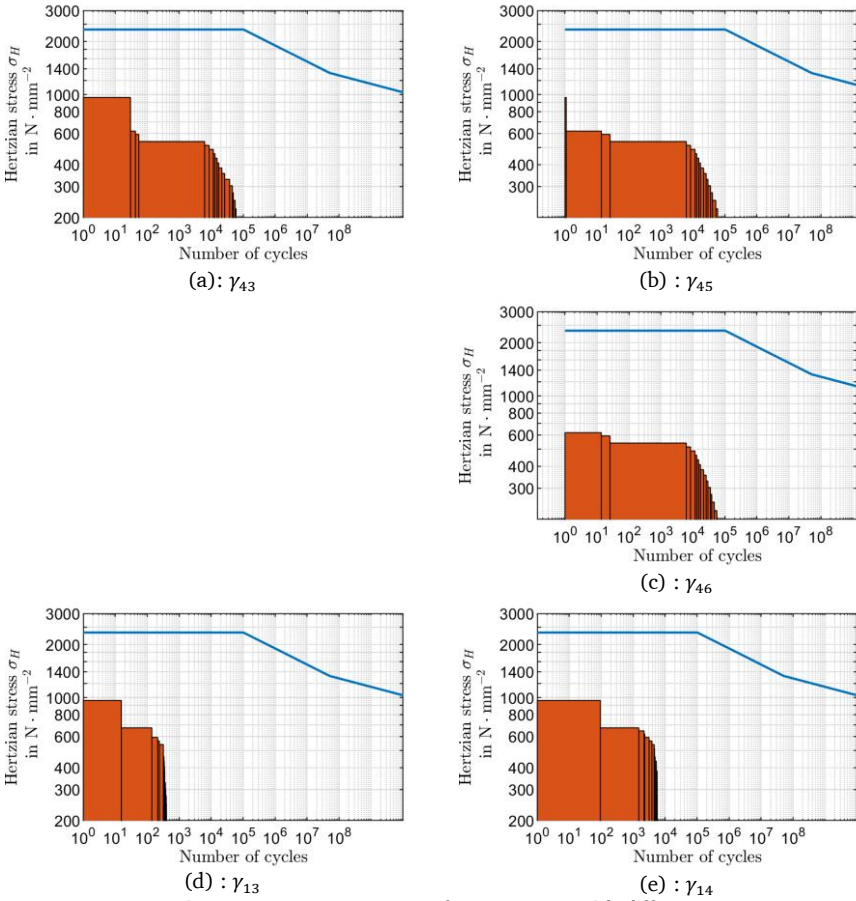


Figure 8.36: Stress spectra of G2.51.P neg with different γ .

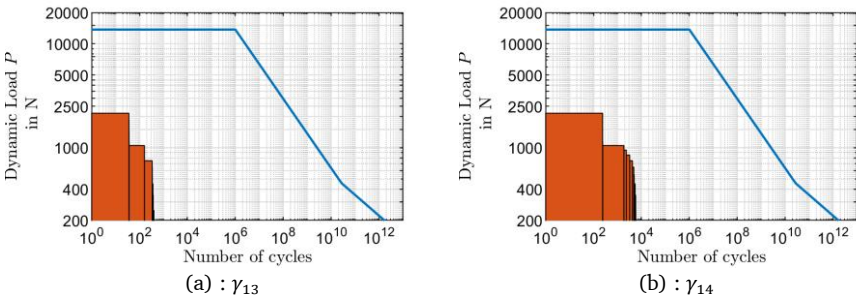


Figure 8.37: Load spectra of B2.1.3 with different γ .

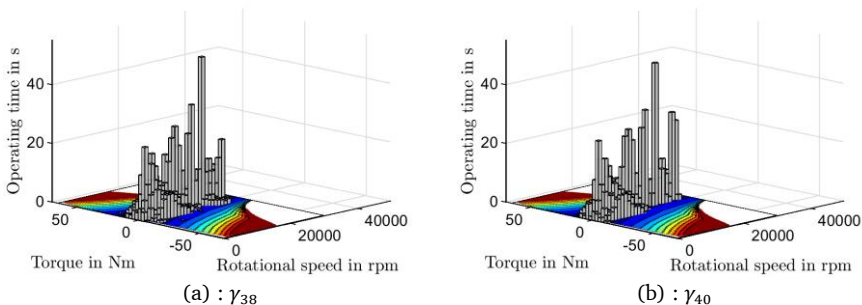


Figure 8.38: Operating duration of the EM1 with different γ .

8.4.4 Evaluation Summary

At γ_1 , the parts in the SPT1 fatigue more severely, since it is mostly the EM1 that propels the vehicle, and the ST2 stays mainly in the neutral gear position. As γ increases, the MCOS seeks lower average accumulated fatigue in the prediction horizon. The MCOS requests, firstly, the ST2 to shift out of the neutral gear position more often. It, secondly, reduces T_1 and increases T_2 . Due to the exponential nature of the fatigue (see Sections 8.1.3, 8.2.1, and 8.3), the average value of $D_{B_{1.1.2}}$ and $D_{B_{2.1.3}}$ decreases, notwithstanding the rise of $D_{B_{2.1.3}}$. The shift count in the WLTC rises until the level near γ_{31} (see Figure 8.33), with which $D_{B_{1.1.2}}$ and $D_{B_{2.1.3}}$ are closest to each other in Figure 8.25.

As γ rises further, $D_{B_{2.1.3}}$ is larger than $D_{B_{1.1.2}}$ (see Figure 8.25). The MCOS further reduces the average value of the accumulated fatigue primarily through reducing $D_{B_{2.1.3}}$, which is sensitive to the torque spikes in the shift processes (see Section 8.3.2). Therefore, less shifts are requested and the shift count in Figure 8.33 decreases after γ_{31} . At the same time, the torques of the EM1 and the EM2 are arranged in such a way that $D_{B_{1.1.2}}$, $D_{B_{2.1.3}}$, and their average value decrease (see the discussion of the load spectra change from the state with γ_{31} through γ_{37} to γ_{55} in Figure 8.31).

8.5 Summary

This chapter explores MCOSs considering both energy consumption and powertrain service life. It addresses the issue that EMSs can cause more heavily loaded parts to be the service life bottleneck of a powertrain. In some unwanted cases, the value of the bottleneck can be several orders of magnitude smaller than the service lives of other parts.

The MCOS in question requires a service life estimation model. With the requirements for online MCOSs and the characteristics of electric powertrains in mind, the subjects of the service life estimation are narrowed down to bearing fatigue failure, gear tooth root breakage and gear tooth

flank pitting. The method for calculating accumulated fatigue is determined. Under their respective assumptions, bearing fatigue caused by positive and negative torques is simplified to be accumulated linearly; gear fatigue of tooth flank pitting caused by positive and negative torques is separately accumulated; and gear fatigue of tooth root breakage caused by positive and negative torques is accumulated with a conservative factor 1.0. The simulative evaluation on Speed4E powertrain in Section 8.1.4 supports the statement in [36, p. 239] that tooth root breakage in modern vehicle transmission is reduced to the minimal⁶³ through modern design processes, and pitting limit is critical. Therefore, tooth root breakage is excluded from the MCOS.

The formulation of the multi-criteria functional in the MCOS is cautiously discussed. The procedure is, arguably, necessary and recommended for all optimal control problems and optimization problems, especially those in application. The fatigue cost functional is decided to be the average accumulated fatigue as an endpoint functional. The differential equation system of the multi-criteria HOCF is shown to be reducible in such a way that the original formulation can be carried over with only modification of the instantaneous Hamiltonian.

The algorithms developed in Chapter 5 are applied as the solver of the MCOS, which shows the versatility of the algorithms.

Speed4E powertrain is used as an example for the formulation of the MCOS as well as its evaluation. As the fatigue coefficient in the cost functional increases, the MCOS has both SPTs share the loads more evenly. As a result, the powertrain service life extends, while the vehicle consumes more energy. The load spectra and the operating duration of both EMs are investigated to interpret and reason of the results. Interestingly, the MCOS requests firstly more shifts and later reduces shifts during the WLTC, as the fatigue coefficient increases. Due to the exponential nature in the mathematical expression of the fatigue phenomena, the MCOS have both EMs avoid the situation that either one provides a major part of the output torque.

The rise of the shift count is, frankly, undesirable but inevitable in the case of Speed4E powertrain. The predictive EMS has the ST2 mostly in its neutral gear position. As the fatigue coefficient rises yet not large enough to make $\gamma \cdot J_{\text{fatigue}}$ overweigh $\int l dt$ in (8.17) by orders of magnitude, the MCOS has the ST2 shifts out of the neutral gear position at a certain portion of the time so that the fatigue cost functional is reduced and the energy consumption is not totally sacrificed. The larger the fatigue coefficient, the larger the portion of the time. If the MCOS is applied on other powertrains, the shift count does not necessarily rise. It depends on powertrain parameters, e.g. EM parameters, gear ratios, parameters of gears and bearings.

The MCOS can be applied with an adaptive fatigue coefficient. For the one with the fatigue cost

⁶³ The original reference used the word “eliminate”.

functional for bearings, the adaptive fatigue coefficient can be formulated as

$$\log \gamma = \text{exponent}_0 + \text{exponent}_{\text{slope}} \left(\log(\max\{0, D_{B,SPT1}^{\text{flag}}\}) \right. \\ \left. + \log(\max\{0, D_{B,SPT2}^{-\text{flag}}\}) \right), \quad (8.47)$$

$$\text{flag} = \begin{cases} 1, & D_{B,SPT1} < D_{B,SPT2}, \\ -1, & D_{B,SPT1} \geq D_{B,SPT2}. \end{cases} \quad (8.48)$$

They tell that γ grows exponentially as both fatigue distance from each other. exponent_0 is the exponent when $D_{B,SPT1}$ and $D_{B,SPT2}$ are identical. $\text{exponent}_{\text{slope}}$ describes how fast γ changes as the difference between both fatigue changes. $D_{B,SPT1}$ and $D_{B,SPT2}$ can be evaluated every thousand kilometres or hundred days. The results of energy consumption and powertrain service life of such an adaptive MCOS moves along the curve in Figure 8.26, as γ changes. exponent_0 and $\text{exponent}_{\text{slope}}$ can be parameterized for some performance measure with methods such as Bayesian Optimization.

9 Conclusion and Outlooks

Before concluding the results, it is worth recapping the background and the main contents.

BEVs are the main solution to the future EU regulation of zero emission passenger cars. It has been a focus of both academia and industry to extend their range. One of the main approaches is to reduce their energy consumption. Recent studies have shown that the two-drive topology and the multi-speed topology help to do so. It is natural to combine both concepts and to design a two-drive multi-speed topology for BEVs. Due to its more than one DoF, an EMS is necessary to exploit the potential of such a topology for reducing total energy consumption. On the other hand, a multi-speed topology brings a discrete variable, i.e. gear positions. Shift processes are involved in EMSs. They are mostly assumed neglectable. [25] and [33] have pointed out that such a practice can cause frequent shifting.

This thesis is initiated by developing an online EMS for a two-drive multi-speed BEV (Speed4E). With its continuous and discrete dynamics in mind, to design an EMS is categorized as an HOCP. For the purpose of online operation, an MPC framework is adopted. Speed4E powertrain is formulated as a hybrid system that considers its shift processes. The HMP is applied and the algorithms for the whole solution process are developed. The predictive EMS embedded with the HMP takes shape. The minimization of the instantaneous Hamiltonian is mathematically examined, and Lemmas 1, 2 and 3 are produced. The time complexity of the algorithms is analysed. In the WLTC driving cycle simulation, the predictive EMS outperforms the EMS naive optimal w.r.t. energy consumption, reduced by 0.26 %, and shift count, reduced by 63.41 %. Due to its online characteristic, the predictive EMS falls short of the EMS global optimal mildly. The durations of the driving cycle simulations support the results of the time complexity analysis.

Through managing power flows, EMSs influence load profiles of parts and components in a powertrain and, therefore, their service lives. As the service life of a powertrain is limited by its part with the shortest service life, it is undesirable to have its parts and components loaded in such way that their service lives is several orders of magnitude apart.

A MCOS is developed in this thesis. Through more evenly dividing loads between the SPT1 and the SPT2, the MCOS extends the powertrain service life. It carries over the same algorithms developed for the predictive EMS embedded with the HMP solution. Tooth flank pitting and bearing fatigue are considered in the service life estimation model. The bottleneck of Speed4E powertrain w.r.t. service life in the WLTC is B1.1.2. With an increasing fatigue coefficient, the MCOS extends the service lives of all parts in the SPT1—B1.1.2 belongs to the SPT1—and, therefore, the service life of Speed4E powertrain. On the other hand, the service lives of all parts in the SPT2 reduced, and the vehicle consumes more energy.

Multi-drive multi-speed electric powertrains are natural further development of two-drive multi-

speed electric powertrains. It seems incomplete to stop with developing EMSs for the latter.

The hybrid system formulation of Speed4E powertrain is generalized to accommodate multi-drive multi-speed electric powertrains. A general multi-drive multi-speed electric powertrain is proposed. The hybrid system formulation for different operation complexity and the corresponding HOCPs are presented.

The detailed discussion of the results can be found in the summary sections in their corresponding chapters. In what follows, a summary and conclusions on the contribution of this thesis as well as an outlook on future works are provided.

Conclusion


This thesis explores EMSs for two-drive multi-speed BEVs “vertically” and “horizontally”.

The “vertical” results are those relate to the two most important attributes of an online EMS: 1) how good can its performance measure be, i.e. how low can energy consumption be? 2) how fast can it compute? They are developed and evaluated with Speed4E BEV as an example.

The popular assumption of neglectable shift processes is examined. It is discarded in this thesis based on two reasons: 1) An EMS normally generates controls of torques and gear positions with a frequency of at least 1 Hz. The shift durations of most common transmissions in today’s market are too long to be ignored; 2) The EMSs with the assumption cause so high shift count that impair drivability (see Sections 1.1 and 3.3 for the reasons, respectively). Through considering shift processes, the hybrid system model from this thesis captures the dynamics of a powertrain more properly. It provides a *new perspective* to model two-drive multi-speed electric powertrains.

The HMP is adopted as the solution method in the predictive EMS for its accuracy and low time complexity. The idea of Bang-Bang controls inspires a *novel* initialization strategy for the solution. It helps the algorithms in this thesis to overcome the shortcoming of IMs in term of initialization, and its time complexity is low. Such an initialization strategy for IMs in the context of EMSs is *unseen in the literature*. The Broyden method is adopted to find the solution of the 2-D TPBVP that is transformed from the HOCP by the HMP. It is worth mentioning that the hybrid system is modelled in such a way that excludes state jumps and, therefore, relieves the solver from the requirement of predefined shift sequences that is found in the literature (see Section 3.2.3). Furthermore, the algorithms for the whole solution process are suitable for multidimensional problems, while most works on EMSs in the literature focused on 1-D problems and their methods are difficult to be extrapolated into multidimensional problems, if possible at all (see Sections 3.2.1.2 and 3.2.2.1). As a result, it holds promises for applying the solver to other HOCPs with the same setting.

A more proper model and an accurate solution method in the MPC framework contribute to the



first attribute. Compared to an optimal EMS that ignores shift processes, i.e. the EMS naive optimal, the predictive EMS reduces energy consumption both in the WLTC and in a real-world driving data based driving cycle simulations. At the same time, the predictive EMS reduces the shift counts significantly in both cases. A predictive EMS embedded with the HMP is *novel*.

The hybrid system model together with the shift processes has increased the complexity of the HOCP and, frankly, raised the time complexity of the algorithms. To mitigate the effect, the continuous control function during the shift processes are modelled mainly time dependent. Furthermore, the Bang-Bang controls inspired initialization and the Broyden method contribute to the second attribute. They as a whole outperform SQP w.r.t. the iteration number for solving TPBVPs. SQP is chosen as the bench mark for its wide availability. Most importantly, the minimization of power consumption is mathematically analysed. With the prerequisite of convex power losses w.r.t. torques, the minimization is proved to be equivalent to a convex optimization in a standard form. Additionally, the optimization domain is proved to be reducible. The Lemmas produced by the proofs reduce the time complexity of the algorithm of the predictive EMS. They can be transformed to the proofs and the treatments for the optimization problems with similar settings in other applications and studies, even those that are not relevant to EMSs. These Lemmas are, as far as the author knows, *unseen in the literature*.

The convexity of the power losses w.r.t. torques might not hold, e.g. when either angular velocity or torque is close to zero. One alternative to apply the Lemmas is to separate the optimization domain. By doing so, only the minimization in a small optimization domain needs to be solved as an NLP problem. Furthermore, if the nonconvexity only occurs in a small enough domain, e.g. when the torque is between -0.01 Nm and 0.01 Nm, and the minimal increment of the torque set points in application is beyond that range, the nonconvexity can be ignored.

As mentioned in Section 3.2.2.1 and verified in Section 6.3.4, an accurate vehicle speed predictor helps a predictive EMS to reduce energy consumption. Developing and training a deep NN model for the purpose of speed prediction has gained traction in academia. This thesis chooses to use a statistical mode for that purpose and to focus on other components in a predictive EMS. The reason is twofold: 1) The model and the solution method relate more directly to EMSs. A proper model benefits all EMSs for two-drive multi-speed BEVs. The solution method, the developed algorithms and the mathematical analysis benefit all problems that can be formulated with a similar setting as the HOCP in the thesis; 2) Theoretical analysis and empirical rigor of contemporary machine learning, especially deep learning, is a matter of debate in recent years [177]. It is out of the scope of this thesis to develop a state-of-the-art NN model *systematically* to predict vehicle speed.

The “horizontal” results are those relate to broadening the scope of the thesis in two aspects: 1) to generalize and to extend the hybrid system formulation and the EMS developed in this thesis;

2) to incorporate a neglected perspective, i.e. powertrain service life in a MCOS.

Based on the hybrid system of Speed4E powertrain, the general multi-drive multi-speed electric powertrain is modelled *for the first time*. Specifically, the general multi-drive powertrain has M EMs, $M \geq 2$, $M \in \mathbb{N}$. Each EM is connected with either a single-speed or a multi-speed ST with the same topology as the ST2 of Speed4E powertrain. The model does not cover all multi-speed transmission topologies. However, the hybrid automaton (the model structure) and the switching conditions among sub-automata—a sub-automaton represents a ST—remain, as long as the discrete states representing shift processes are modified according to other types of shift processes. One has to change the sub-automata, if n -speed STs are considered, $n \geq 3$, $n \in \mathbb{N}$.

The HOCP for the general multi-drive multi-speed powertrain is proposed *for the first time*, which can be solved by the algorithms developed for the predictive EMS with the optimality conditions by the HMP provided in Section 7.2.


The hybrid system and the HOCP of the general multi-drive multi-speed powertrain provides a *unified* framework for modelling and designing EMSs for a multi-drive multi-speed powertrain. It is noteworthy that the complexity of the general multi-drive multi-speed powertrain is encoded in the hybrid system, the HOCP and the optimality condition mainly through the admissible discrete control sets and the sample space of the combination of the discrete state of the whole system. It is showcased, how to achieve different operation complexity through modifying them. This underlines the adjective “*unified*”. It is not mandatory to apply the algorithms developed in this thesis. Notwithstanding the time complexity growing exponentially with M , the DP can solve the proposed HOCP.

Due to the lack of powertrain data, the general multi-drive multi-speed powertrain is not evaluated in simulation environment in this thesis.

With the previous work [37] in mind, this thesis is the *second* to propose a MCOS considering powertrain service life. With the exponential nature of fatigue w.r.t. torques in mind, the MCOS can extend the powertrain service life by several times, but sacrifices energy consumption by at most less than one quarter, compared to the predictive EMS.

A service life estimation model that follows DIN 3990 and DIN ISO 281 is developed for the MCOS. The simulation study in Section 8.1.4 shows that, for BEVs, loads and fatigue caused by positive and negative torques are comparable. Therefore, operational fatigue strength under alternating loads is essential to service life estimation of electric powertrains, while it is suggested to be neglectable in the case of powertrains of convectional vehicles (see Section 8.1.3.1). Fatigue caused by alternating loads is modelled under the assumptions summarized in Section 8.5. With more refined models, fatigue caused by alternating loads can be more accurately determined.

The influence of the torques and the shift processes on the service life is investigated in Section



8.3. Due to the exponential nature of fatigue rate w.r.t. torques, the torque spikes during the shift processes cause the relevant parts to fatigue noticeably. A similar situation can be expected, if one evaluates a powertrain in which some parts withstand a high load for a short period of time in shift processes or other transient processes.

The multi-criteria HOCF is systematically formulated, for which different formulations are evaluated. This procedure is nontrivial. EMSs with whatever solution methods stress numerical convergence, which can be compromised by extra coefficients and functionals/functions introduced by multi-criteria cost functionals/functions. For better convergence, the fatigue cost functional is decided to be in the form of an endpoint functional, unlike the common practice of MCOSs that adds a second cost function multiplying a coefficient. It is worth mentioning that formulation and its benefits are individual to problems. It is not recommended to incorporate a second cost functional in the form of an endpoint functional blindly but to follow the same train of thought for evaluation and analysis.

Outlooks

The predictive EMS can be improved by a more accurate vehicle speed predictor. It can be achieved through not only deploying an advanced NN model but also incorporating holistic information into the inputs of the predictor, such as route information and information from smart city internet of things networks. It is worth researching, how sensitively prediction accuracy influences the performance of a predictive EMS. Furthermore, studies on efforts for building a vehicle speed predictor vs. its benefit to the performance in both quantitative and qualitative ways can be helpful for outlining future research focuses of predictive EMSs.

At the current stage, the HMP under the category IMs is identified as the suitable solution method for the predictive EMS. There is always room for improvement in convergence rate and accuracy. One can improve the algorithms listed in Chapter 5. Other solution methods can be more suitable when their relevant algorithms evolve. One can also research on other solution methods and produce Lemmas under which other solution methods compute with lower time complexity for the purpose of EMSs, just like Lemmas 1,2 and 3 in this thesis.

To avoiding frequent shifting, it is a common practice in engineering to define a minimal wait time between shifts. One can include this aspect in a comparative study.

The predictive EMS with the HMP solution method can be flashed into a control unit. Hardware-in-the-Loop can be used to verify the simulation study in this thesis, which provides value in engineering.

Fatigue of electric powertrains can be more accurately calculated with the help of experiments and more accurate modelling methods, especially those for determining fatigue caused by alternating loads. EM thermal lifetime can be incorporated in the service life estimation model with the

help of a model of the closed loop dynamics of the thermal management system. On the other hand, if lower accuracy can be tolerated, one can assume that the temperatures of both EMs remain unchanged in a short prediction horizon. Such an assumption is not sound at all for offline EMSs solving a multi-criteria HOCP with a lengthy speed profile.

An adaptive MCOS like the one proposed in Section 8.5 is worth researching.

The predictive EMS and the MCOS developed in this thesis can be integrated in design processes for future powertrains. Design parameters of EMs and STs can be optimized iteratively based on driving cycle simulations. The results of the MCOS can be considered during construction dimensioning.

Studies on ownership costs of powertrains can be performed to evaluate the benefit of the service life extension. Construction dimensioning can even be less conservative, if the MCOS helps to extend the powertrain service life to an acceptable extent. Such an approach is less attractive for passenger cars, for their consumers are sensitive to energy consumption. It might be attractive for heavy-duty vehicles, whose owners value ownership costs more, if the approach is proved to reduce ownership costs.

A unified framework to model the general multi-drive multi-speed powertrain and to design its EMS provides plenty of room for research. It can be used to build an automatic optimization tool for electric powertrain topology optimization and component design. It can be used to build an automatic EMS generation tool.

Appendix A The Parameters of Speed4E Powertrain and the vehicle

The reduced rotational inertia about the rotational axis of the wheels are

$$I_{\text{red}} = I_{\text{red,PT1}} + I_{\text{red,PT2}} + 4 \cdot I_{\text{Wheel}}. \quad (\text{A.1})$$

The rotational inertia of the sub-powertrain 1 (PT1), EM1 and the ST1 combined, about the rotational axis of the wheels can be calculated by

$$I_{\text{red,PT1}} = (I_{\text{EM1}} + I_{\text{Sun}} + I_{\text{InputShaft}}) \cdot i_1^2 + (I_{\text{Carrier}} + I_{\text{SplineShaft}} + I_{\text{PinoinFD}}) \cdot i_{1,2}^2 \cdot i_{FD}^2. \quad (\text{A.2})$$

The rotational inertia of the sub-powertrain 2 (PT2), EM2 and the ST2 combined, about the rotational axis of the wheels is gear position dependent, which can be calculated by

$$I_{\text{red,PT2}} = \begin{cases} (I_{\text{EM2}} + I_{\text{PinionS1}} + I_{\text{InputShaft}}) \cdot i_2^2(G) \\ + \left(I_{\text{WheelS1}} + I_{\text{PinionS2G1}} + I_{\text{PinionS2G2}} + \frac{I_{\text{WheelS2G2}}}{i_{2,2}^2} + I_{\text{CounterShaft}} \right) \cdot \frac{i_2^2(G)}{i_{2,1}^2} \\ + (I_{\text{WheelS2G1}} + I_{\text{PinionS3}} + I_{\text{SS}} + I_{\text{OutputShaft}}) \cdot i_{2,3}^2 \cdot i_{FD}^2, & G = 1, \\ \\ (I_{\text{EM2}} + I_{\text{PinionS1}} + I_{\text{InputShaft}}) \cdot i_2^2(G) \\ + \left(I_{\text{WheelS1}} + I_{\text{PinionS2G1}} + I_{\text{PinionS2G2}} + \frac{I_{\text{WheelS2G2}}}{i_{2,2}^2} + I_{\text{CounterShaft}} \right) \cdot \frac{i_2^2(G)}{i_{2,2}^2} \\ + (I_{\text{WheelS2G1}} + I_{\text{PinionS3}} + I_{\text{SS}} + I_{\text{OutputShaft}}) \cdot i_{2,3}^2 \cdot i_{FD}^2, & G = 2, \\ \\ (I_{\text{PinionS3}} + I_{\text{SS}} + I_{\text{OutputShaft}}) \cdot i_{2,3}^2 \cdot i_{FD}^2, & G = 3. \end{cases} \quad (\text{A.3})$$

The rotational inertias of the components in Speed4E powertrain are listed in Table A. 1.

Table A. 1: Rotational inertia of the components in Speed4E powertrain

| Component | Symbol | Value | Unit |
|--|--------------------|------------------------|------------------------------|
| Rotor of the EM1 | I_{EM1} | 371.1×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Sun gear of the ST1 | I_{Sun} | 9×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Input shaft (EM1 to sun gear) | $I_{InputShaft}$ | 25×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Carrier of ST1 | $I_{Carrier}$ | 2700×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Spline shaft (carrier to FD) | $I_{SplineShaft}$ | 3730×10^{-6} | m/s^2 |
| Pinion of the gear set with the ratio $i_{1,2}$ | $I_{PinoinFD}$ | 480×10^{-6} | kg/m^3 |
| Rotor of the EM2 | I_{EM2} | 371.1×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Pinion of gear set stage 1 | $I_{PinionS1}$ | 35×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Wheel of gear set stage 1 | $I_{WheelS1}$ | 14660×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Counter shaft | $I_{CounterShaft}$ | 134×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Pinion of gear set stage 2 in 1 st gear | $I_{PinionS2G1}$ | 418×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Wheel of gear set stage 2 in 1 st gear | $I_{WheelS2G1}$ | 67500×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Pinion of gear set stage 2 in 2 nd gear | $I_{PinionS2G2}$ | 3065×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Wheel of gear set stage 2 in 2 nd gear | $I_{WheelS2G2}$ | 25600×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Output shaft | $I_{OutputShaft}$ | 1052×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| shift sleeve | I_{SS} | 15×10^{-6} | $\text{kg} \cdot \text{m}^2$ |
| Wheel | I_{Wheel} | 4.2 | $\text{kg} \cdot \text{m}^2$ |

The parameters for calculation of longitudinal dynamics are listed in the following table.

Table A. 2: Parameters of longitudinal dynamics

| Parameter | Symbol | Value | Unit |
|-------------------------|--------------|-----------------------------|-------------------|
| Vehicle mass | m | 1320 | kg |
| Fronted area | A_f | 2.38 | m ² |
| Aerodynamic coefficient | c_{aero} | 0.294 | - |
| Rolling coefficient | c_{roll} | Velocity dependent function | - |
| Acceleration of gravity | g | 9.8 | m/s ² |
| Density of air | ρ_{air} | 1.2 | kg/m ³ |

The vehicle parameters are provided by the BMW Group. The parameters of the parts in the powertrain are provided by the Gear Research Center (FZG) from Technical University Munich and the Institute for Drive Systems and Power Electronics from Leibniz University Hannover.

Appendix B Speed4E Powertrain Power Losses

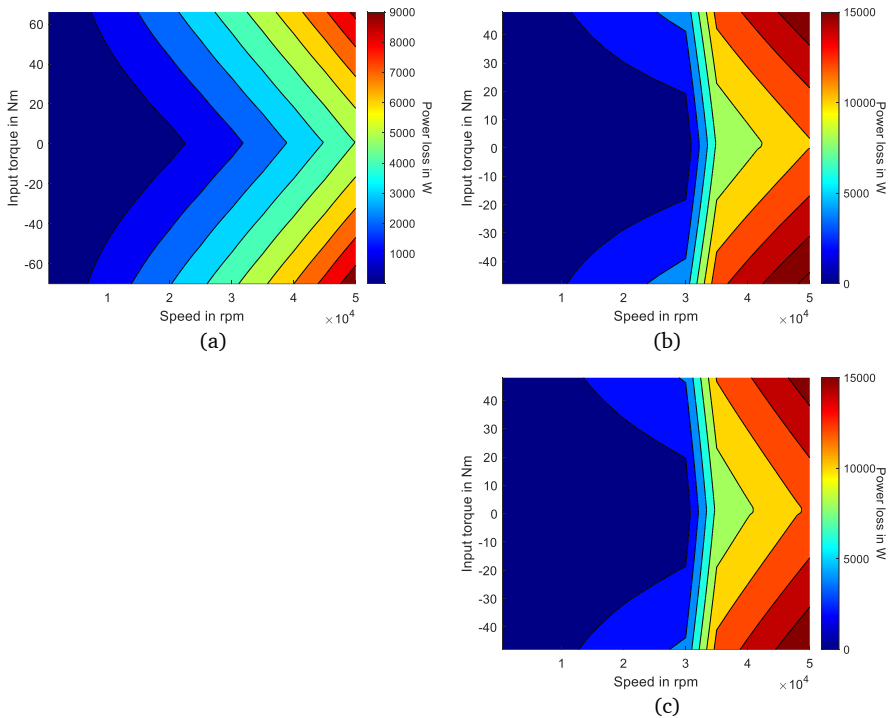


Figure B. 1: Power loss maps of (a) the ST1, (b) the ST2 in the 1st gear position and (c) the ST2 in the 2nd gear position.

The power losses of both STs, shown in Figure B. 1, are provided by the Gear Research Center, Technical University Munich. The power losses of both EMs, shown in Figure B. 2, are provided by the Institute for Drive Systems and Power Electronics, Leibniz University. The results of both LEs, shown in Figure B. 2, are provided by both the Institute for Drive Systems and Power Electronics, Leibniz University Hannover and the Lenze SE. The maximal and minimal torques are marked by black curves.

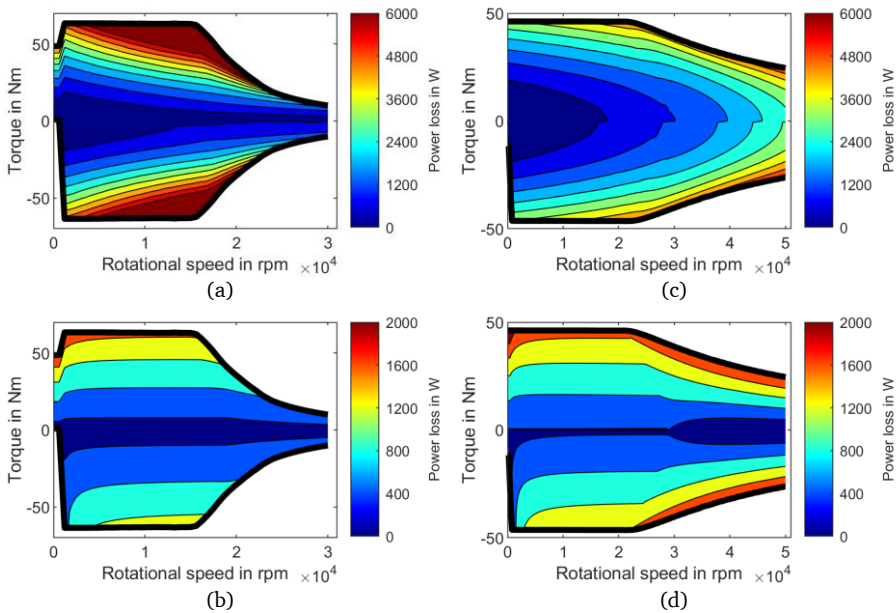


Figure B. 2: Power loss maps of (a) the EM1, (b) the LE1, (c) the EM2 and (d) the LE2.

Appendix C Dynamics of a M-Drive Multi-Speed Powertrain

Longitudinal Dynamics

Suppose a BEV is propelled by a powertrain with M EMs that are connect to M STs. A discrete set Its longitudinal acceleration can be determined by

$$\dot{v} = b(G, v, T_{\text{total}}) = \frac{\frac{T_{\text{total}}}{r} - c_{\text{roll}}(v) \cdot \text{mass} \cdot g - 0.5 \cdot c_a \cdot \rho_{\text{air}} \cdot A_f \cdot v^2}{\text{mass}_{\text{eq}}(G)}. \quad (\text{C.1})$$

The total output torque is provided by all EMs

$$T_{\text{total}} = \sum_{n \in \Pi} T_{\text{ST}n} \cdot i_n, \quad (\text{C.2})$$

$$T_{\text{ST}n} = \phi_n(\omega_n, T_n) = T_n - \frac{PL_{\text{ST}n}(\omega_n, T_n)}{\omega_n}, \quad (\text{C.3})$$

$$\Pi := \{I, II, \dots, M\}. \quad (\text{C.4})$$

The derivative of the angular velocity over itself of an EM can be determined by

$$\frac{\partial \dot{\omega}_n}{\partial \omega_n} = \frac{\partial b}{\partial \omega_n}(G, v, T_{\text{total}}) \quad (\text{C.5})$$

Note that v is related to ω_n , and, with (C.8) and (C.9) in mind, T_{total} is related to ω_n . There is

$$\begin{aligned} \frac{\partial \dot{\omega}_n}{\partial \omega_n} &= \frac{\partial b}{\partial \omega_n}(G, v, T_{\text{total}}) \\ &= \frac{\partial v}{\partial \omega_n} \cdot \frac{\partial b}{\partial v}(G, v, T_{\text{total}}) + \frac{\partial \phi_n}{\partial \omega_n} \cdot \frac{\partial T_{\text{total}}}{\partial \phi_n} \cdot \frac{\partial b}{\partial T_{\text{total}}}(G, v, T_{\text{total}}) \\ &= -\frac{i_n}{r m_{\text{eq}}(G_n)} \cdot \left(\frac{i_n}{r \omega_n^2} \frac{\partial PL_{\text{ST}n}}{\partial \omega_n}(\omega_n, u_n) + \frac{r}{i_n} \left(\frac{dc_r}{dv}(v) \cdot m \cdot g + c_a \rho A_f v \right) \right). \end{aligned} \quad (\text{C.6})$$

Necessary modifications to model the hybrid system with the first assumption Section 7.1.1

It is assumed that *one ST shifts at a time*.

An additional discrete variable $n \in \Pi, \Pi := \{I, II, \dots, M\}$, is introduced to index the sub-automata. Let G_n represents the gear position of the ST to which the n -th sub-automaton corresponds. The function to calculate the effective torque for each ST is ϕ_{n,G_n} . The function to calculate the input torque based on the effective torque is ϕ'_{n,G_n} . The gear ratio of the n -th ST at G_n gear position is $i_n(G_n)$. The angular velocities and the torques of the EMs form the continuous states $\mathbf{x} := (\omega_I, \omega_{II}, \dots, \omega_M)^T$ and the continuous controls $(T_I, T_{II}, \dots, T_M)^T$, respectively.

The reduced inertia J_{red} in (4.5) is dependent on the gear position of each ST. They are encoded in a combination $\mathbf{G} := (G_I, G_{II}, \dots, G_M)$.

In the π sub-automaton, a discrete set $\mathcal{Q}_n := \{q_{n,1}, q_{n,2}, \dots, q_{n,9}, q_{n,4,a}, q_{n,4,b}, \dots, q_{n,9,d}\}$ includes the discrete states and the sub-states of the sub-automaton that are defined in the same way as in Section 4.4. w_n denotes the discrete control in the n -th sub-automaton. Table 4.2 and Table 4.3 are taken over to formulate \mathcal{B}_n , the admissible discrete control set in the π sub-automaton, in which q is replaced by q_n . The manifolds M_n are defined in the same way as in Section 4.4.2. The dynamics of the n -th sub-automaton are defined in the same way as in Section 4.4.1 and Section 4.4.2.

Discrete states $q_{n,1}$, $q_{n,2}$ and $q_{n,3}$

All STs stay in their fixed gear positions. Therefore, there is $\forall n \in \Pi, q_n = G_n$. The indexed system dynamics can be derived in a similar way to (4.14) from (4.6) as

$$\begin{aligned} \dot{\mathbf{x}}_q &= \mathbf{f}_q(\mathbf{x}, T_{\text{total}}) \\ &= \frac{b(\mathbf{G}, v, T_{\text{total}})}{r} \cdot (i_{G_I}, i_{G_{II}}, \dots, i_{G_M})^T, \quad \forall t \in [t_{0,q_n}, t_{f,q_n}], \end{aligned} \quad (\text{C.7})$$

$$\mathbf{G} = (G_I, G_{II}, \dots, G_M), \quad (\text{C.8})$$

$$T_{\text{total}} = \sum_{n \in \Pi} i_{G_n} \cdot \phi_{n,G_n}(\omega_n, T_n), \quad (\text{C.9})$$

$$v = \frac{\omega_I}{i_{G_I}} \cdot r, \quad (\text{C.10})$$

$$\mathbf{x}(t_{0,q_n}) = \mathbf{x}_{0,q_n}, \quad (\text{C.11})$$

where t_{0,q_n} and t_{f,q_n} denote the start and end time of q_n ; the initial state \mathbf{x}_{0,q_n} is given.

Discrete states $q_{n,4}$, $q_{n,5}$, \dots , $q_{n,9}$: the n -th ST performs a shift

“Balance” and “Decrease” phases

As discussed in Section 4.3, the duration of the “Balance” phase is the smaller one between the duration for T_n to reach zero ($\tau_{n,\text{Fade}}$) and the duration for the rest to reach its maximum

$(\tau_{\Pi \setminus n, \text{Fade}}) \cdot \setminus$ denotes “exclude”. They are determined by

$$\tau_{\Pi \setminus n, \text{Fade}} = \frac{1}{TS_{\text{Bal}}} \cdot (T_{\text{outMax}, \Pi \setminus n}(t_{0, \pi, j}) - T_{\text{out}, \Pi \setminus n}(t_{0, \pi, j})), \quad (\text{C. 12})$$

$$T_{\text{outMax}, \Pi \setminus n}(t) = \sum_{n' \in \Pi \setminus n} i_{n'}(q_{n'}) \cdot \phi_{n', G_{n'}}(\omega_{n'}(t), T_{n', \text{max}}(\omega_{n'})), \quad (\text{C. 13})$$

$$T_{\text{out}, \Pi \setminus n}(t) = \sum_{n' \in \Pi \setminus n} i_{n'}(q_{n'}) \cdot \phi_{n', G_{n'}}(\omega_{n'}(t), T_{n'}(t)), \quad (\text{C. 14})$$

$$\tau_{n, \text{Fade}} = \frac{(T_n(t_{0, q_n})) \cdot i_n(q_n(t_{0, n, j}^-))}{TS_{\text{Bal}}}, \quad (\text{C. 15})$$

$$\tau_{\text{Bal}} = \min(\tau_{\Pi \setminus n, \text{Fade}}, \tau_{n, \text{Fade}}), \quad (\text{C. 16})$$

where $T_{\text{out}, \Pi \setminus n}$ denotes the output torque provided by all EMs excluding the n -th EM. $T_{\text{outMax}, \Pi \setminus n}$

denotes the maximum value of $T_{\text{out}, \Pi \setminus n}$. By (C. 12),

(C.13) and (C.14), all EMs excluding the n -th EM ramp up their torques in such a way that they reach their maximal torques at the same time. Since all STs other than the n ST are in fixed-gear position, given the assumption that one ST shifts at a time, $i_{n'}(q_{n'})$ that appears in the following texts is the gear ratio of the n' -th ST. TS_{Bal} denotes the rate of the change of output torque on the wheels in the “Balance” phase; $t_{0, n, j}$ the time point when q_n enters $q_{n, j}$.

A “Decrease” phase takes place, if $\exists n' \in \Pi \setminus n$, $T_{n'}$ reaches its maximum before T_n reaches zero. The duration of the “Decrease” phase is

$$\tau_{\text{Decr}} = \begin{cases} \frac{T_n(t_{0, n, j} + \tau_{\text{Bal}}) \cdot i_n(q_n(t_{0, n, j}^-))}{TS_{\text{Decr}}}, & \tau_{\text{Bal}} < \tau_{n, \text{Fade}}, \\ 0, & \tau_{\text{Bal}} = \tau_{n, \text{Fade}} \end{cases} \quad (\text{C. 17})$$

$$\tau_a = \tau_{\text{Bal}} + \tau_{\text{Decr}}, \quad (\text{C. 18})$$

where TS_{Decr} denotes the slope of the change of output torque on the wheels in the “Decrease” phase. If the duration of the “Balance” phase equals $\tau_{n, \text{Fade}}$, the “Decrease” phase does not take place, which means τ_{Decr} equals zero. The duration of the sub-state $q_{n, j, a}$ (τ_a) is the sum of the duration of the “Balance” and “Decrease” phases. In this sub-state, the torque variables in \mathbf{u} are determined by

$$T_{\text{out},\Pi \setminus n} = \begin{cases} T_{\text{out},\Pi \setminus n}(t_{0,q_n}) + TS_{\text{Bal}} \cdot (t - t_{0,n,j}), & t \in (t_{0,n,j}, t_{0,n,j} + \tau_{\text{Bal}}], \\ T_{\text{outMax},\Pi \setminus n}(t_{0,n,j} + \tau_{\text{Bal}}), & t \in (t_{0,n,j} + \tau_{\text{Bal}}, t_{0,n,j} + \tau_a], \end{cases}$$

$$TS_{\text{Bal}} = \sum_{n' \in \Pi \setminus n} TS_{\text{Bal},n'} \quad (\text{C.19})$$

$$TS_{\text{Bal},n'} = \frac{i_{n'}(q_{n'}) \cdot [T_{n',\text{max}}(\omega_{n'}) - \phi_{n',G_{n'}}(\omega_{n'}(t_{0,n,j}), T_{n'}(t_{0,n,j}))]}{T_{\text{outMax},\Pi \setminus n}(t_{0,n,j}) - T_{\text{out},\Pi \setminus n}(t_{0,n,j})} \cdot TS_{\text{Bal}}, \quad n' \in \Pi \setminus n$$

$$T_{n'} = \begin{cases} T_{n'}(t_{0,q_n}) + \frac{TS_{\text{Bal},n'}}{i_{n'}(q_{n'})} \cdot (t - t_{0,n,j}), & t \in (t_{0,n,j}, t_{0,n,j} + \tau_{\text{Bal}}], \\ T_{n',\text{max}}(\omega_{n'}), & t \in (t_{0,n,j} + \tau_{\text{Bal}}, t_{0,n,j} + \tau_a], \end{cases} \quad (\text{C.21})$$

$$T_n = \begin{cases} \phi'_{n,G_n} \left(\omega_n, \frac{T_{\text{total}} - T_{\text{out},\Pi \setminus n}(t)}{i_n(q(t_{0,n,j}^-))} \right), & t \in (t_{0,n,j}, t_{0,n,j} + \tau_{\text{Bal}}], \\ T_n(t_{0,n,j} + \tau_{\text{Bal}}) - \frac{TS_{\text{Decr}}}{i_n(q(t_{0,n,j}^-))} \cdot (t - t_{0,n,j} - \tau_{\text{Bal}}), & t \in (t_{0,n,j} + \tau_{\text{Bal}}, t_{0,n,j} + \tau_a], \end{cases} \quad (\text{C.22})$$

where $T_{n'}$, one of the torque of the EMs other than n -th EM, fades to its maximum with a slope dependent on $TS_{\text{Bal},n'}$, while T_n fades to zero firstly in accordance to the rising of the other EMs and secondly dependent on TS_{Decr} , if necessary. $TS_{\text{Bal},n'}$ is a part of TS_{Bal} , where the proportion is dependent on the deviation of $T_{n'}$ to $T_{n',\text{max}}$. Since the DC is not disengaged, (C.7)-(C.9) hold with

$$\mathbf{G} = (G_I, G_{II}, \dots, G_M), \quad G_n = q(t_{0,n,j}^-), \quad t \in (t_{0,n,j}, t_{0,n,j} + \tau_a]. \quad (\text{C.23})$$

An autonomous transition from state $q_{n,j,a}$ to $q_{n,j,b}$ happens, when the time-dependent switching manifold $M_{ab}(t) = 0$, which is defined as

$$M_{ab}(t) = t - t_{0,n,j} - \tau_a. \quad (\text{C.24})$$

First ‘‘Synchronize’’ and ‘‘Disengage’’ phases

In $q_{\pi,j,b}$, the first ‘‘Synchronize’’ phase takes place, followed by the ‘‘Disengage’’ phase. Both phases are assumed to last for constant time, as discussed in Section 4.3. The duration of this sub-state (τ_b) and the torques are

$$\begin{aligned}\tau_b &= \tau_{\text{syn1}} + \tau_{\text{Disengage}} & (\text{C.25}) \\ T_{\text{out},\Pi/n} &= T_{\text{total}} - \phi_{n,q}(t_{0,n,j}^-)(\omega_n, 0) \cdot i_n \left(q_n(t_{0,n,j}^-) \right), \quad t \in (t_{0,n,j,b}, t_{0,n,j,b} + \tau_b], & (\text{C.26}) \\ T_{\text{total}} &= p(\mathbf{G}, v, \dot{v}) & (\text{C.27}) \\ T_n &= 0, \quad t \in (t_{0,n,j,b}, t_{0,n,j,b} + \tau_b], & (\text{C.28})\end{aligned}$$

where the rest EMs propel the vehicle, also compensating the drag torque of ST. The requested output torque is determined by $p(\cdot)$ based on (4.6), in which \mathbf{G} determines I_{red} . The value of each torque remains to be determined by the EMS, which can be considered an optimization problem. $t_{0,\pi,j,b}$ is the time point when the sub-state $q_{\pi,j,b}$ is activated. Since the DC is not disengaged, (C.7)-(C.9) hold with

$$\mathbf{G} = (G_I, G_{II}, \dots, G_M), G_n = q(t_{0,n,j}^-), \quad t \in (t_{0,n,j,b}, t_{0,n,j,b} + \tau_b]. \quad (\text{C.29})$$

An autonomous transition from sub-state $q_{n,j,b}$ to $q_{n,j,c}$ happens, when the time-dependent switching manifold $M_{bc}(t) = 0$, which is defined as

$$M_{bc}(t) = t - t_{0,n,j,b} - \tau_b. \quad (\text{C.30})$$

Second ‘‘Synchronize’’ and ‘‘Engage’’ phases

In $q_{\pi,j,c}$, the second ‘‘Synchronize’’ phase with a duration τ_{syn2} takes place, followed by the ‘‘Engage’’ phase with constant duration, as discussed in Section 4.3. τ_{syn2} depends on the vehicle speed and the shift process $q_{n,j}$. It includes a) the duration for EM2 to accelerate τ_{acc} , negative acceleration in this case, and b) the duration of the APC τ_{APC} . The duration of this sub-state (τ_c) and the torques are

$$\tau_c = \tau_{\text{acc}} + \tau_{\text{APC}} + \tau_{\text{Engage}} = \tau_{\text{syn2},q_{n,j}}(v) + \tau_{\text{Engage}} \quad (\text{C.31})$$

$$T_{\text{out},\Pi/n} = \min(T_{\text{outMax},\Pi/n}, T_{\text{total}}), \quad t \in (t_{0,n,j,c}, t_{0,n,j,c} + \tau_c], \quad (\text{C.32})$$

$$T_{\text{total}} = p(\mathbf{G}, v, \dot{v}), \quad (\text{C.33})$$

$$T_n = 0, \quad t \in (t_{0,n,j,c}, t_{0,n,j,c} + \tau_c], \quad (\text{C.34})$$

$$G_n = 3, \quad t \in (t_{0,n,j,c}, t_{0,n,j,c} + \tau_c], \quad (\text{C.35})$$

where the gear position is neutral, since the DC is disengaged. The dynamics follow

$$\dot{x}_{n'} = \frac{b(\mathbf{G}, v, T_{\text{out}, \Pi \setminus n})}{r} \cdot i_{n'}(q_{n'}), \quad n' \in \Pi \setminus n, \quad \forall t \in (t_{0,n,j,c}, t_{0,n,j,c} + \tau_c], \quad (\text{C.36})$$

$$\dot{x}_n = \frac{T_2}{I_{n,\text{input}}}, \quad \forall t \in (t_{0,n,j,c}, t_{0,n,j,c} + \tau_c], \quad (\text{C.37})$$

$$v = \frac{\omega_{n'}}{i_{n'}(q_{n'})} \cdot r, \quad n' \in \Pi \setminus n, \quad (\text{C.38})$$

$$T_n = \begin{cases} T_{n,\text{min}}, & t \in (t_{0,n,j,c}, t_{0,n,j,c} + \tau_{\text{acc}}], \\ 0, & t \in (t_{0,n,j,c} + \tau_{\text{acc}}, t_{0,n,j,c} + \tau_c] \end{cases}, \quad (\text{C.39})$$

$$\tau_{\text{acc}} = \frac{\left(\frac{v}{r} \cdot i_n(\overline{\omega}_{q_{n,j,d}}) - \omega_n(t_{0,n,j,c})\right) I_{\text{ST2,input}}}{T_n}, \quad (\text{C.40})$$

where $I_{\pi,\text{input}}$ is the rotational inertia around π EM rotor axis including the rotor of π EM, the first stage and the second stage of the π ST. For $q_{n,j} \in \{q_{n,4}, q_{n,5}, \dots, q_{n,9}\}$, $\overline{\omega}_{q_{n,j,d}}$ is equivalent to the target gear position. All torques other than T_n remain to be determined by the EMS.

An autonomous transition from the sub-state $q_{n,j,c}$ to $q_{n,j,d}$ happens, when the time-dependent switching manifold $M_{cd}(t) = 0$, which is defined as

$$M_{cd}(t) = t - t_{0,n,j,c} - \tau_c. \quad (\text{C.41})$$

“Increase” and second “Balance” phases

In $q_{n,j,d}$, $T_{\text{out}, \Pi \setminus n}$, $T_{n'}$ and T_n fade to $T_{\text{outNew}, \Pi \setminus n}$, $T_{n',\text{new}}$ and $T_{n,\text{new}}$, respectively, which are determined by the EMS for the new gear position. Similar to the sub-state $q_{n,j,a}$, both “Balance” and “Increase” phases might take place. Their durations are determined by

$$\tau_{\Pi \setminus n, \text{Fade}} = \frac{T_{\text{out}, \Pi \setminus n}(t_{0,\pi,j,d}) - T_{\text{outNew}, \Pi \setminus n}}{TS_{\text{Bal}}}, \quad (\text{C.42})$$

$$\tau_{n, \text{Fade}} = \frac{(T_{n,\text{new}} - T_n(t_{0,n,j,d})) \cdot i_n(\overline{\omega}_{q_{n,j}})}{TS_{\text{Bal}}}, \quad (\text{C.43})$$

$$\tau_{\text{Bal}} = \min(\tau_{\Pi \setminus n, \text{Fade}}, \tau_{n, \text{Fade}}), \quad (\text{C.44})$$

$$\tau_{\text{Incr}} = \begin{cases} \frac{(T_{n,\text{new}} - T_n(t_{0,n,j,d} + \tau_{\text{Bal}})) \cdot i_n(\overline{\omega}_{q_{n,j,d}})}{TS_{\text{Incr}}}, & \tau_{\text{Bal}} < \tau_{n, \text{Fade}}, \\ 0, & \tau_{\text{Bal}} = \tau_{n, \text{Fade}}, \end{cases} \quad (\text{C.45})$$

$$\tau_d = \tau_{\text{Bal}} + \tau_{\text{Incr}}, \quad (\text{C.46})$$

where $\overline{\omega}_{q_{n,j,d}}$ is equivalent to the target gear, which is defined by the autonomous admissible control set. In this sub-state, the torques are determined by

$$TS_{\text{Bal},n'} = \frac{i_{n'}(q_{n'}) \cdot [T_{n',\text{new}} - \phi_{n',G_{n'}}(\omega_{n'}(t_{0,n,j}), T_{n'}(t_{0,n,j}))]}{T_{\text{outNew},\Pi \setminus n} - T_{\text{out},\Pi \setminus n}(t_{0,n,j,d})} \cdot TS_{\text{Bal}}, \quad n' \in \Pi \setminus n \quad (\text{C.47})$$

$$T_{n'} = \begin{cases} T_{n'}(t_{0,n,j,d}) + \frac{TS_{\text{Bal},n'}}{i_{n'}(q_{n'})} \cdot (t - t_{0,n,j,d}), & t \in (t_{0,n,j,d}, t_{0,n,j,d} + \tau_{\text{Bal}}], \\ T_{n',\text{new}}, & t \in (t_{0,n,j,d} + \tau_{\text{Bal}}, t_{0,n,j,d} + \tau_d], \end{cases} \quad (\text{C.48})$$

$$T_n = \begin{cases} \phi'_{n,G_n} \left(\omega_n, \frac{T_{\text{total}} - T_{\text{out},\Pi \setminus n}(t)}{i_n(\varpi_{q_{n,j,d}})} \right), & t \in (t_{0,n,j,d}, t_{0,n,j,d} + \tau_{\text{Bal}}], \\ T_n(t_{0,n,j,d} + \tau_{\text{Bal}}) + \frac{TS_{\text{Incr}}}{i_n(\varpi_{q_{n,j,d}})} \cdot (t - t_{0,n,j,d} - \tau_{\text{Bal}}), & t \in (t_{0,n,j,d} + \tau_{\text{Bal}}, t_{0,n,j,d} + \tau_d]. \end{cases} \quad (\text{C.49})$$

Since the DC of the new gear position is engaged, (C.7)-(C.9) hold with

$$G = \varpi_{q_{n,j,d}}, \quad t \in (t_{0,n,j,d}, t_{0,n,j,d} + \tau_d]. \quad (\text{C.50})$$

An autonomous transition from the sub-state $q_{n,j,d}$ the next discrete state $\varpi_{q_{n,j,d}}$ happens, when the time-dependent switching manifold $M_{j\varpi_{q_{n,j,d}}}(t) = 0$, which is defined as

$$M_{j\varpi_{q_{n,j,d}}}(t) = t - t_{0,n,j,d} - \tau_d. \quad (\text{C.51})$$

Shift process in other categories

In the case of $T_{\text{total}} \leq 0$, replace $T_{n',\text{max}}$ with $T_{n',\text{min}}$ in (C.12), (C.13), (C.20) and (C.21). $T_{\text{outMax},\Pi \setminus n}$ is consequently $T_{\text{outMin},\Pi \setminus n}$. In the case of $\omega_{n,\text{target}} > \omega_n$, replace $T_{n,\text{min}}$ with $T_{n,\text{max}}$ in (C.39).

Necessary modifications to model the hybrid system with the second assumption in Section 7.1.2

It is assumed that *at most* $M - 1$ STs can shift at the same time.

Further restriction is applied to the system: the controlled switching of all sub-automata take place at the same time. More specifically, at a time point t , $\forall n \in \Pi$, $\varpi_n(t)$ can be different to $\varpi_n(t^-)$, if

and only if $q_n(t^-) \leq 3$.

At a time point t , the indexes of the sub-automata that are in a shift process form a set Π_{shift} . Those that remain in their fixed gear positions form a set $\Pi \setminus \Pi_{\text{shift}}$. The first strategy to control the shift processes is to have each phase in a shift process of all sub-automata in Π_{shift} finish at the same time. For a sub-automaton π that enters a shift process, the dynamics and the predefined control in (C.7)-(C.51) hold but with the following changes:

- The set $\Pi \setminus n$ is replaced by $\Pi \setminus \Pi_{\text{shift}}$.
- Since all the phases in the shift processes are aligned, the time point during a shift process can be expressed by the time variable of any sub-automaton.
- T_n is not directly determined. An example in “Balance” and “Decrease” phases, i.e. subscript “a”, is shown:

The duration was calculated by (C.12)-(C.18). They are changed to

$$\tau_{\Pi \setminus \Pi_{\text{shift}}, \text{Fade}} = \frac{1}{TS_{\text{Bal}}} \cdot (T_{\text{outMax} \Pi \setminus \Pi_{\text{shift}}}(t_{0,n,j}) - T_{\text{out}, \Pi \setminus \Pi_{\text{shift}}}(t_{0,n,j})), \quad (\text{C.52})$$

$$T_{\text{outMax}, \Pi \setminus \Pi_{\text{shift}}}(t) = \sum_{n' \in \Pi \setminus n} i_{n'}(q_{n'}) \cdot \phi_{n', G_{n'}}(\omega_{n'}(t), T_{n', \text{max}}(\omega_{n'})), \quad (\text{C.53})$$

$$T_{\text{out}, \Pi \setminus \Pi_{\text{shift}}}(t) = \sum_{n' \in \Pi \setminus n} i_{n'}(q_{n'}) \cdot \phi_{n', G_{n'}}(\omega_{n'}(t), T_{n'}(t)), \quad (\text{C.54})$$

$$\tau_{\Pi_{\text{shift}}, \text{Fade}} = \frac{1}{TS_{\text{Bal}}} \cdot (T_{\text{out}, \Pi_{\text{shift}}}(t_{0,n,j}^-) - T_{\text{outZero}, \Pi_{\text{shift}}}(t_{0,n,j}^-)), \quad (\text{C.55})$$

$$T_{\text{outZero}, \Pi_{\text{shift}}}(t) = \sum_{n \in \Pi_{\text{shift}}} i_n(q_n(t)) \cdot \phi_{n, G_n}(\omega_n(t), 0), \quad (\text{C.56})$$

$$T_{\text{out}, \Pi_{\text{shift}}}(t) = \sum_{n \in \Pi_{\text{shift}}} i_n(q_n(t)) \cdot \phi_{n, G_n}(\omega_n(t), T_n(t)) \quad (\text{C.57})$$

$$\tau_{\text{Bal}} = \min(\tau_{\Pi \setminus \Pi_{\text{shift}}, \text{Fade}}, \tau_{\Pi_{\text{shift}}, \text{Fade}}), \quad (\text{C.58})$$

$$\tau_{\text{Decr}} = \begin{cases} \frac{T_{\text{out}, \Pi_{\text{shift}}}(t_{0,n,j} + \tau_{\text{Bal}})}{TS_{\text{Decr}}}, & \tau_{\text{Bal}} < \tau_{n, \text{Fade}}, \\ 0, & \tau_{\text{Bal}} = \tau_{n, \text{Fade}} \end{cases} \quad (\text{C.59})$$

$$\tau_a = \tau_{\text{Bal}} + \tau_{\text{Decr}}, \quad (\text{C.60})$$

The torques $T_{n'}, n' \in \Pi \setminus \Pi_{\text{shift}}$ are determined in the same way as in (C.19)-(C.22). The torques $T_n, n \in \Pi_{\text{shift}}$ are determined

$$\begin{aligned}
& T_{\text{out},\Pi_{\text{shift}}} \\
&= \begin{cases} T_{\text{total}} - T_{\text{out},\Pi \setminus \Pi_{\text{shift}}}(t), t \in (t_{0,n,j}, t_{0,n,j} + \tau_{\text{Bal}}], \\ T_{\text{out},\Pi_{\text{shift}}}(t_{0,n,j} + \tau_{\text{Bal}}) - TS_{\text{Decr}} \cdot (t - t_{0,n,j} - \tau_{\text{Bal}}), t \in (t_{0,n,j} + \tau_{\text{Bal}}, t_{0,n,j} + \tau_a], \end{cases} \quad (\text{C.61})
\end{aligned}$$

$$TFactor_{\text{Bal},n} = \frac{i_n(q_n(t_{0,n,j}^-)) \cdot \phi_{n,G_n}(\omega_n(t_{0,n,j}), T_n(t_{0,n,j}))}{T_{\text{out},\Pi_{\text{shift}}}(t_{0,n,j}^-) - T_{\text{outZero},\Pi_{\text{shift}}}(t_{0,n,j}^-)}, \quad n \in \Pi_{\text{shift}} \quad (\text{C.62})$$

$$T_n = \phi'_{n,G_n}(\omega_n, TFactor_{\text{Bal},n} \cdot T_{\text{out},\Pi_{\text{shift}}}), \quad (\text{C.63})$$

where $TFactor_{\text{Bal},n}$ is a factor that considers the deviation between T_n and 0. By doing so, the sub-states $q_{n,j,a}$, $n \in \Pi_{\text{shift}}$, $j \in \{4,5,6,7\}$ of all sub-automaton in Π_{shift} end at the same time.

- The duration of the second ‘‘Synchronize’’ and ‘‘Engage’’ phases, i.e. subscript ‘‘c’’, is the highest value of the all sub-automata in Π_{shift} :

$$\tau_c = \max\{\tau_{\text{syn}2,q_{n,j}}\} + \tau_{\text{Engage}}, \quad n \in \Pi_{\text{shift}}. \quad (\text{C.64})$$

The strategy to control the shift processes can also be changed. For instance, T_n , $n \in \Pi_{\text{shift}}$, changes individually with a factor. (C.61) remains. (C.62) and (C.63) are replaced by the equations that have T_n changes with a slope of $\frac{T_{\text{SBal}}}{N_{\Pi_{\text{shift}}}}$, where $N_{\Pi_{\text{shift}}}$ is the size of Π_{shift} . The duration of ‘‘Synchronize’’ and ‘‘Engage’’ phases is individually decided, after which T_n changes with individual slope. $T_{\text{out},\Pi \setminus \Pi_{\text{shift}}}$ compensates the change of $T_{\text{out},\Pi_{\text{shift}}}$. The formulation is omitted.

Appendix D Cost Function Minimization

Minimization of the cost function is formulated as

$$\min_{\mathbf{u}_q} l(\mathbf{x}, \mathbf{u}_q), \quad (\text{D.1})$$

$$l(\mathbf{x}, \mathbf{u}) = \mathbf{x}^T \cdot \mathbf{u} + PL_{EM1}(\omega_1, u_1) + PL_{EM2}(\omega_2, u_2), \quad (\text{5.7})$$

subject to

$$u_1 \in [T_{1,\min}(\omega_1), T_{1,\max}(\omega_1)], \quad (\text{5.13})$$

$$u_2 \in [T_{2,\min}(\omega_2), T_{2,\max}(\omega_2)], \quad (\text{5.14})$$

$$\dot{\mathbf{x}}_q = \mathbf{f}_q(\mathbf{x}, T_{\text{total}}) \quad (\text{4.14})$$

$$= \frac{b(q, v, T_{\text{total}})}{r} \cdot \begin{pmatrix} i_1 \\ i_2(q) \end{pmatrix}, \quad (\text{4.15})$$

$$T_{\text{total}} = \phi_1(\omega_1, u_1) \cdot i_1 + \phi_{2,q}(\omega_2, u_2) \cdot i_2(q) \quad (\text{4.16})$$

$$q \in \{1, 2\}, \quad (*)$$

where in (*) only q_1 and q_2 are considered, since the torques are determined by (4.7)-(4.9) in neutral gear position and the continuous controls are predefined during a shift process ($q \in \mathcal{Q} \setminus \{q_1, q_2, q_3\}$). The minimization is subject to the inequality constraints (5.8) and (5.9) as well as the equality constraint (4.16). (5.15) and (5.16) are not considered, since \mathbf{x} and q are given.

Equivalent convex optimization

Lemma 1: the problem (D.1) is equivalent to a convex optimization problem in a standard form.

Proof of Lemma 1:

Firstly, the power losses are examined. For studies of EMS, power losses are often considered quadratic polynomial with two variables (angular velocity and torque), in which the parameters of the 2nd degree terms are positive [26]. In such settings, the power loss is convex. In this work, the power losses are modelled as look-up tables based on the simulation results. The power losses of both EMs and both STs at different angular velocities are illustrated in Figure D. 1, which are a clearer view of the contour plots Figure B. 1: in Appendix B. By observation, PL_{ST} and PL_{EM} are convex functions in the direction of torque T .

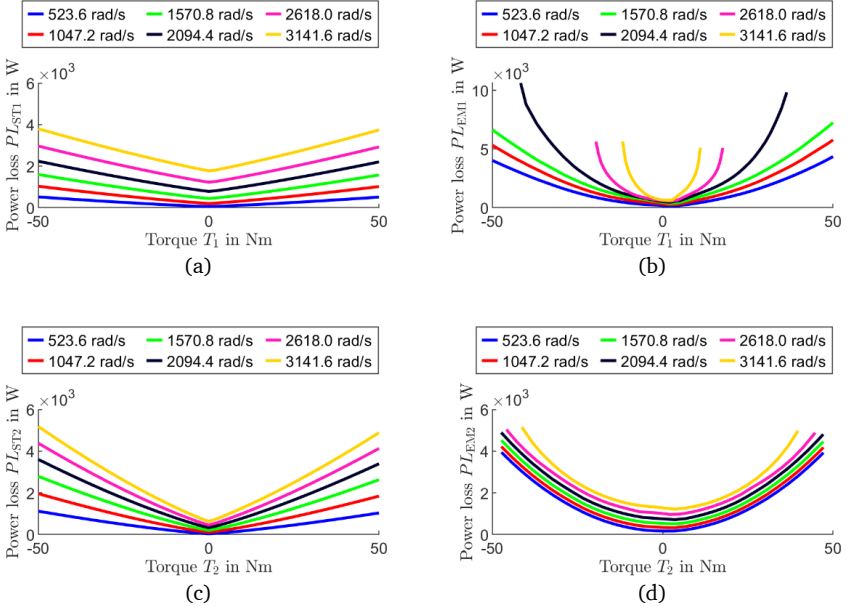


Figure D. 1: Power loss at different angular velocities of (a) ST1, (b) EM1, (c) ST2 and (d) EM2

Secondly, change of variables of the optimization problem:

The function $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}$ in (4.16) is more rigidly defined as

$$T_{ST} = \phi(\omega, T) = T - \frac{PL_{ST}(\omega, T)}{\omega}, T \in \mathcal{T}(\omega), \quad (\text{D.2})$$

$$\mathcal{T}(\omega) = \{T \in \mathbb{R} \mid T_{min}(\omega) \leq T \leq T_{max}(\omega)\}. \quad (\text{D.3})$$

Since the angular velocity ω is given for the optimization problem, the function is rewritten as $\phi_\omega: \mathbb{R} \rightarrow \mathbb{R}$

$$T_{ST} = \phi_\omega(T) = T - \frac{PL_{ST}(\omega, T)}{\omega}, T \in \mathcal{T}_\omega, \quad (\text{D.4})$$

$$\mathcal{T}_\omega = \{T \in \mathbb{R} \mid T_{min}(\omega) \leq T \leq T_{max}(\omega)\}, \quad (\text{D.5})$$

where domain \mathcal{T}_ω is bounded by the maximum and the minimum torque of the EM depending on its angular velocity. \mathcal{T}_ω , a continuous 1-D line, is a convex set. In this domain, ϕ_ω is concave, as

PL_{ST} is convex. By definition of concavity, $\forall c \in [0,1], \forall T_a, T_b \in \mathcal{T}_\omega$, the following holds

$$\phi_\omega(c \cdot T_a + (1 - c) \cdot T_b) \geq c \cdot \phi_\omega(T_a) + (1 - c) \cdot \phi_\omega(T_b). \quad (\text{D.6})$$

Figure D. 2(a) and (b) show the output of ϕ_ω of ST1 and ST2at first gear, i.e. $\phi_{\omega,1}$ and $\phi_{\omega,2,q_1}$, at different ω , which can be observed to be monotonic. The inverse function ϕ_ω^{-1} therefore exists. The domain of ϕ_ω^{-1} is the range of ϕ_ω

$$\text{dom}\phi_\omega^{-1} = \{\phi_\omega(T) \mid T \in \mathcal{T}_\omega\}, \quad (\text{D.7})$$

which is continuous and bounded by $\phi_\omega(\sup \mathcal{T}(\omega))$ and $\phi_\omega(\inf \mathcal{T}(\omega))$ on \mathbb{R} . “sup” and “inf” represent supremum and infimum. (D.7) is a convex set.

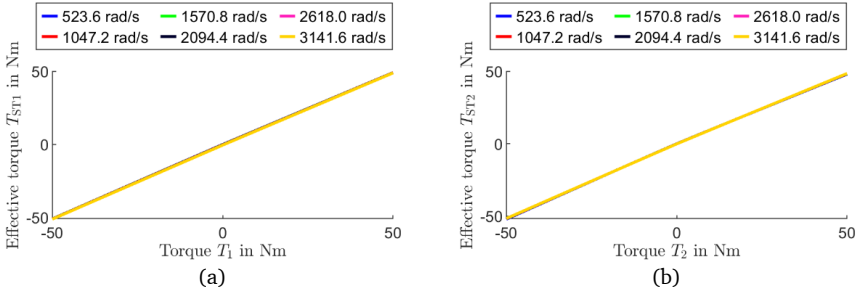


Figure D. 2: Effective torques of the ST1 (a) and the ST2 in 1st gear (b) at different angular velocities

$\forall c \in [0,1], \forall T_m, T_n \in [\phi_\omega(\inf \mathcal{T}(\omega)), \phi_\omega(\sup \mathcal{T}(\omega))]$, two real numbers α and β are defined as

$$\alpha = \phi_\omega^{-1}(c \cdot T_m + (1 - c) \cdot T_n), \quad (\text{D.8})$$

$$\beta = c \cdot \phi_\omega^{-1}(T_m) + (1 - c) \cdot \phi_\omega^{-1}(T_n), \quad (\text{D.9})$$

which still belong to $\mathcal{T}(\omega)$. Place α in ϕ_ω

$$\begin{aligned} \phi_\omega(\alpha) &= c \cdot T_m + (1 - c) \cdot T_n \\ &= c \cdot \phi_\omega(\phi_\omega^{-1}(T_m)) + (1 - c) \cdot \phi_\omega(\phi_\omega^{-1}(T_n)). \end{aligned} \quad (\text{D.10})$$

By the inequality from (D.6)

$$\phi_\omega(\alpha) \leq \phi_\omega(c \cdot \phi_\omega^{-1}(T_m) + (1 - c) \cdot \phi_\omega^{-1}(T_n)) = \phi_\omega(\beta). \quad (\text{D.11})$$

Since ϕ strictly increases, there is $\alpha \leq \beta$, i.e.

$$\phi_{\omega}^{-1}(c \cdot T_m + (1 - c) \cdot T_n) \leq c \cdot \phi_{\omega}^{-1}(T_m) + (1 - c) \cdot \phi_{\omega}^{-1}(T_n), \quad (\text{D.12})$$

$\forall c \in [0,1], \forall T_m, T_n \in \mathcal{J}_{ST}(\omega)$. ϕ_{ω}^{-1} is therefore a convex function.

Use $\phi_{\omega}^{-1}(T_{ST})$ to substitute T_1 and T_2 ($\mathbf{u} := (T_1, T_2)$) in (5.7) and the cost function is transformed to

$$\begin{aligned} l_{\mathbf{x},q} \left(\phi_{1,\omega_1}^{-1}(T_{ST1}), \phi_{2,\omega_2,q}^{-1}(T_{ST2}) \right) \\ = \omega_1 \cdot \phi_{1,\omega_1}^{-1}(T_{ST1}) + \omega_2 \cdot \phi_{2,\omega_2,q}^{-1}(T_{ST2}) \\ + PL_{EM1} \left(\omega_1, \phi_{1,\omega_1}^{-1}(T_{ST1}) \right) + PL_{EM2} \left(\omega_2, \phi_{2,\omega_2,q}^{-1}(T_{ST2}) \right), \end{aligned} \quad (\text{D.13})$$

whose minimization is

$$\min_{(T_{ST1}, T_{ST2})^T} l_{\mathbf{x},q} \left(\phi_{1,\omega_1}^{-1}(T_{ST1}), \phi_{2,\omega_2,q}^{-1}(T_{ST2}) \right), \quad (\text{D.14})$$

subject to

$$T_{ST1} \in \{ \phi_{ST1,\omega_1}(T) \mid T \in \mathcal{J}_{EM1,\omega_1} \}, \quad (\text{D.15})$$

$$T_{ST2} \in \{ \phi_{ST2,\omega_2}(T) \mid T \in \mathcal{J}_{EM2,\omega_2} \}, \quad (\text{D.16})$$

$$\dot{\mathbf{x}}_q = \mathbf{f}_q(\mathbf{x}, T_{\text{total}}) \quad (\text{D.17})$$

$$= \frac{b(q, v, T_{\text{total}})}{r} \cdot \begin{pmatrix} i_1 \\ i_2(q) \end{pmatrix}, \quad (\text{D.18})$$

$$T_{\text{total}} = T_{ST1} \cdot i_1 + T_{ST2} \cdot i_2(q), \quad (4.7)$$

$$q \in \{1,2\}, \quad (*)$$

where \mathbf{x} and q are subscripted in (D.13) and (D.14) to avoid confusion, since the cost function is optimized w.r.t. torques. The last two terms in (D.13) are convex functions, since they are function composition of convex functions. (D.13) as a whole is a convex function, since it's a sum of convex functions with non-negative weights. Same as the reasoning for (D.7), (D.15) and (D.16) are convex sets. (4.7) is an affine function. Therefore, (D.14) is a convex optimization in standard form [63, pp. 136-137].

Lastly, the equivalency. The control \mathbf{u}_q in (D.1) and the effective torques $(T_{ST1}, T_{ST2})^T$ are one-to-one, since ϕ_{ST1,ω_1} and ϕ_{ST2,ω_2} are monotonic. If $(T_{ST1}^*, T_{ST2}^*)^T$ solves (D.14), $(\phi_{ST1,\omega_1}^{-1}(T_{ST1}^*), \phi_{ST2,\omega_2}^{-1}(T_{ST2}^*))^T$ solves (D.1). The original problem (D.1) is therefore equivalent to the problem (D.14) by change of variables [63, pp. 130-131].

■

Remark: To obtain the function $\phi'(\omega, T_{ST})$ introduced at (4.7)-(4.11), firstly perform $\phi_{\omega}^{-1}(T_{ST})$ on the discretized ω , then perform interpolation.

Upper and lower bounds of the torques

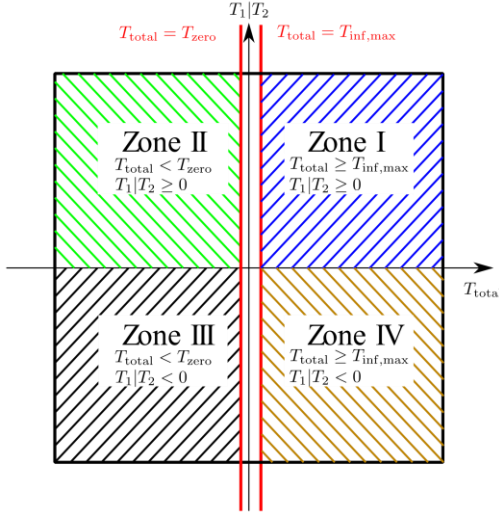


Figure D. 3: Torque range divided into several zones

Lemma 2: Suppose $(\mathbf{T}_{ST1}^*, \mathbf{T}_{ST2}^*)^T$ solves (D.14) and $(\mathbf{u}_1^*, \mathbf{u}_2^*)^T = (\boldsymbol{\phi}_{ST1, \omega_1}^{-1}(\mathbf{T}_{ST1}^*), \boldsymbol{\phi}_{ST2, \omega_2}^{-1}(\mathbf{T}_{ST2}^*))^T$ is the equivalent solution to (D.1). Following inequality holds

$$\begin{cases} u_1^* \in [0, T_{1, \max}(\omega_1)], \\ u_2^* \in [0, T_{2, \max}(\omega_2)], \\ T_{\text{total}} \geq T_{\text{inf, max}}, \end{cases} \quad (D.19)$$

$$\begin{cases} u_1^* \in [T_{1, \min}(\omega_1), 0], \\ u_2^* \in [T_{2, \max}(\omega_2), 0], \\ T_{\text{total}} < T_{\text{zero}}, \end{cases}$$

$$\begin{cases} T_{ST1}^* \in [\boldsymbol{\phi}_{\omega_1, ST1}(0), \boldsymbol{\phi}_{\omega_1, ST1}(T_{1, \max}(\omega_1))], \\ T_{ST2}^* \in [\boldsymbol{\phi}_{\omega_2, ST2, q}(0), \boldsymbol{\phi}_{\omega_2, ST2, q}(T_{2, \max}(\omega_2))], \\ T_{\text{total}} \geq T_{\text{total, inf, max}}, \end{cases} \quad (D.20)$$

$$\begin{cases} T_{ST1}^* \in [\boldsymbol{\phi}_{\omega_1, ST1}(T_{1, \min}(\omega_1)), \boldsymbol{\phi}_{\omega_1, ST1}(0)], \\ T_{ST2}^* \in [\boldsymbol{\phi}_{\omega_2, ST2, q}(T_{2, \max}(\omega_2)), \boldsymbol{\phi}_{\omega_2, ST2, q}(0)], \\ T_{\text{total}} < T_{\text{zero}}. \end{cases}$$

Proof of Lemma 2:

(D.20) is the mapping of (D.19) through $\boldsymbol{\phi}_{ST1, \omega_1}$ and $\boldsymbol{\phi}_{ST2, \omega_2, q}$. It's necessary to only prove (D.19).

$(T_{\text{total}}, u_1^*)$ and $(T_{\text{total}}, u_2^*)$ are not in “Zone II” or “Zone IV” at the same time, since it contradicts the equality constraint (4.16).

If Lemma 2 holds, both points $(T_{\text{total}}, u_1^*)$ and $(T_{\text{total}}, u_2^*)$ fall in either “Zone I” or “Zone III”, which is proved by contradiction in what follows.

Step I: consider the $T_{\text{total}} \geq T_{\text{inf,max}}$. Assume $u_1^* \geq 0$, i.e. “Zone I”, and $u_2^* < 0$, i.e. “Zone IV”. Choose a $u_1 \in [0, u_1^*]$ in such a way that $u_2 = 0$ can be determined by (4.16). Choose a small increment $\delta u_1 > 0$ in such a way that $u_1 + \delta u_1 < u_1^*$. The equality constraint (4.16) determines the new torque of EM2 corresponding to $u_1 + \delta u_1$ and it is denoted as $u_2 - \delta u_2$, $\delta u_2 > 0$. As (4.16) is not affine, $(u_1 + \delta u_1, u_2 - \delta u_2)^T$ does not lie on the line segment between $(u_1, u_2)^T$ and $(u_1^*, u_2^*)^T$.

The functions $\phi_{\omega_1, \text{ST1}}$ and $\phi_{\omega_2, \text{ST2}, q}$ map $(u_1, u_2)^T$, $(u_1 + \delta u_1, u_2 - \delta u_2)^T$ and $(u_1^*, u_2^*)^T$ to $(T_{\text{ST1}}, T_{\text{ST2}})^T$, $(T_{\text{ST1}} + \delta T_{\text{ST1}}, T_{\text{ST2}} - \delta T_{\text{ST2}})^T$ and $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$, respectively, where $\delta T_{\text{ST1}} > 0$ and $\delta T_{\text{ST2}} > 0$. $(T_{\text{ST1}}, T_{\text{ST2}})^T$, $(T_{\text{ST1}} + \delta T_{\text{ST1}}, T_{\text{ST2}} - \delta T_{\text{ST2}})^T$ and $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$ all satisfy the affine equality constraint (4.7), which indicates that $(T_{\text{ST1}} + \delta T_{\text{ST1}}, T_{\text{ST2}} - \delta T_{\text{ST2}})^T$ lies on the line segment between $(T_{\text{ST1}}, T_{\text{ST2}})^T$ and $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$. Since (D.13) is a convex function,

$$l_{x,q}(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}}), \phi_{2,\omega_2,q}^{-1}(T_{\text{ST2}})) \geq l_{x,q}(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}} + \delta T_{\text{ST1}}), \phi_{2,\omega_2,q}^{-1}(T_{\text{ST2}} - \delta T_{\text{ST2}})). \quad (\text{D.21})$$

Substitute $(T_{\text{ST1}}, T_{\text{ST2}})^T$ with $(u_1, u_2)^T$

$$l_{x,q}(u_1, u_2) \geq l_{x,q}(u_1 + \delta u_1, u_2 - \delta u_2). \quad (\text{D.22})$$

Consider the deviation between the left- and right-hand sides of (D.22)

$$l_{x,q}(u_1 + \delta u_1, u_2 - \delta u_2) - l_{x,q}(u_1, u_2) \leq 0, \quad (\text{D.23})$$

$$\begin{aligned} \omega_1 \cdot \delta u_1 - \omega_2 \cdot \delta u_2 + PL_{\text{EM1}}(\omega_1, u_1 + \delta u_1) - PL_{\text{EM1}}(\omega_1, u_1) \\ + PL_{\text{EM2}}(\omega_2, u_2 - \delta u_2) - PL_{\text{EM2}}(\omega_2, u_2) \leq 0, \end{aligned} \quad (\text{D.24})$$

subject to

$$\begin{aligned} T_{\text{total}} &= \phi_{1,\omega_1}(u_1) \cdot i_1 + \phi_{2,\omega_2,q}(u_2) \cdot i_2 \\ &= \phi_{1,\omega_1}(u_1 + \delta u_1) \cdot i_1 + \phi_{2,\omega_2,q}(u_2 - \delta u_2) \cdot i_2. \end{aligned} \quad (\text{D.25})$$

Since $T_{\text{total}} \geq T_{\text{total,inf,max}}$ and $u_2 = 0$, there is $u_1 \geq T_{1,\text{inf}}$, where $\frac{\partial PL_{\text{EM1}}}{\partial u_1}(\omega_1, u_1) > 0$, by observation of Figure D. 1(c). By observation of Figure D. 1(b), $\frac{\partial PL_{\text{EM2}}}{\partial u_2}(\omega_2, u_2) < 0$, for $u_2 = 0$ and $\delta u_2 > 0$. Therefore,

$$PL_{EM1}(\omega_1, u_1 + \delta u_1) - PL_{EM1}(\omega_1, u_1) > 0, \quad u_1 \geq T_{1,\text{inf}} \text{ and } \delta u_1 > 0, \quad (\text{D.26})$$

$$PL_{EM2}(\omega_2, u_2 - \delta u_2) - PL_{EM2}(\omega_2, u_2) > 0, \quad u_2 = 0 \text{ and } \delta u_2 > 0. \quad (\text{D.27})$$

Furthermore, by observation of Figure D. 2(a) and (b), $\frac{\partial PL_{ST1}}{\partial u_1}(\omega_1, u_1) > 0$, for $T_1 \geq T_{1,\text{inf}}$, and $\frac{\partial PL_{ST2}}{\partial u_2}(\omega_2, u_2) < 0$, for $u_2 < 0$. Based on (D.4) and the convexity of the power losses, there are $\frac{d\phi_{1,\omega_1}}{du_1} < 1$, for $u_1 \geq T_{1,\text{inf}} > T_{1,\text{ST1},\text{inf}}$, and $\frac{d\phi_{2,\omega_2,q}}{du_2} > 1$, for $u_2 < 0 < T_{2,\text{ST2},\text{inf}}$ (see the beginning of this proof). Rearrange (D.25) to get

$$\begin{aligned} 0 &= \phi_{1,\omega_1}(u_1 + \delta u_1) - \phi_{1,\omega_1}(u_1) + \frac{i_2(q)}{i_1} \cdot (\phi_{2,\omega_2,q}(u_2 - \delta u_2) - \phi_{2,\omega_2,q}(u_2)) \\ &= \frac{d\phi_{1,\omega_1}(u_1)}{du_1} \cdot \delta u_1 - \frac{i_2(q)}{i_1} \cdot \frac{d\phi_{2,\omega_2,q}(u_2)}{du_2} \cdot \delta u_2 \\ &< \delta u_1 - \frac{i_2(q)}{i_1} \cdot \delta u_2 = \frac{1}{\omega_1} \cdot (\omega_1 \cdot \delta u_1 - \omega_2 \cdot \delta u_2). \end{aligned} \quad (\text{D.28})$$

Considering (D.26)-(D.28), the left-hand side of the inequality (D.24) is greater than zero. This gives a contradiction to (D.24) stemmed from the convexity. Therefore, the situation that $u_1^* \geq 0$ and $u_2^* < 0$ does not exist, if $T_{\text{total}} \geq T_{\text{inf,max}}$.

Step II: still consider the case that $T_{\text{total}} \geq T_{\text{total,inf,max}}$. Assume $u_1^* < 0$, i.e. ‘‘Zone IV’’ and $u_2^* \geq 0$, i.e. ‘‘Zone I’’. Choose $u_1 = 0$ and $0 < u_2 < u_2^*$ can be determined by (4.16). Choose a small increment $\delta u_1 > 0$, so that $u_1^* < u_1 - \delta u_1 < 0$. The equality constraint (4.16) determines the new torque of EM2 corresponding to $u_1 - \delta u_1$ and it is denoted as $u_2 + \delta u_2, \delta u_2 > 0$. As (4.16) is not affine, $(u_1 - \delta u_1, u_2 + \delta u_2)^T$ does not lie on the line segment between $(u_1, u_2)^T$ and $(u_1^*, u_2^*)^T$.

The functions $\phi_{\omega_1, \text{ST1}}$ and $\phi_{\omega_2, \text{ST2}, q}$ map $(u_1, u_2)^T$, $(u_1 - \delta u_1, u_2 + \delta u_2)^T$ and $(u_1^*, u_2^*)^T$ to $(T_{\text{ST1}}, T_{\text{ST2}})^T$, $(T_{\text{ST1}} - \delta T_{\text{ST1}}, T_{\text{ST2}} + \delta T_{\text{ST2}})^T$, where $\delta T_{\text{ST1}} > 0$ and $\delta T_{\text{ST2}} > 0$, and $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$ respectively, which all satisfy (4.7). The point $(T_{\text{ST1}} - \delta T_{\text{ST1}}, T_{\text{ST2}} + \delta T_{\text{ST2}})^T$ lies on the line segment between $(T_{\text{ST1}}, T_{\text{ST2}})^T$ and $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$. Since (D.13) is a convex function,

$$l_{x,q} \left(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}}), \phi_{2,\omega_2,q}^{-1}(T_{\text{ST2}}) \right) \geq l_{x,q} \left(\phi_{1,\omega_1}^{-1}(T_{\text{ST1}} - \delta T_{\text{ST1}}), \phi_{2,\omega_2,q}^{-1}(T_{\text{ST2}} + \delta T_{\text{ST2}}) \right). \quad (\text{D.29})$$

A contradiction can be found by considering the deviation of the left and right sides of (D.29) and following the rest of Step I. Therefore, the situation that $u_1^* < 0$ and $u_2^* \geq 0$ does not exist, if $T_{\text{total}} \geq T_{\text{inf,max}}$. Together with the conclusion from Step I, the following holds

$$\begin{aligned} u_1^* &\in [0, T_{1,\max}(\omega_1)], \\ u_2^* &\in [0, T_{2,\max}(\omega_2)], \quad T_{\text{total}} \geq T_{\text{inf,max}}. \end{aligned} \quad (\text{D.30})$$

Repeat Step I and II to prove Lemma 2 by contradictions for the case that $T_{\text{total}} < T_{\text{zero}}$. ■

The torque of EM1 providing a total output torque of $T_{\text{total,inf,max}}$ is determined by

$$T_{1,\text{inf,max}} = \phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total,inf,max}} - \phi_{2,\omega_2,q}(0) \cdot i_2(q)}{i_1} \right), \quad (\text{D.31})$$

while that of EM2 is similarly determined by

$$T_{2,\text{inf,max}} = \phi_{2,\omega_2,q}^{-1} \left(\frac{T_{\text{total,inf,max}} - \phi_{1,\omega_1}(0) \cdot i_1}{i_2(q)} \right). \quad (\text{D.32})$$

Lemma 3: Suppose $(T_{\text{ST1}}^*, T_{\text{ST2}}^*)^T$ solves (D.14) and $(u_1^*, u_2^*)^T = (\phi_{1,\omega_1}^{-1}(T_{\text{ST1}}^*), \phi_{2,\omega_2}^{-1}(T_{\text{ST2}}^*))^T$ is the equivalent solution to (D.1). Following inequality holds

$$\begin{aligned} u_1^* &\in \left[\phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1} \right), T_{1,\text{inf,max}} \right], \\ u_2^* &\in \left[\phi_{2,\omega_2,q}^{-1} \left(\frac{T_{\text{total}} - \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1}{i_2(q)} \right), T_{2,\text{inf,max}} \right], \end{aligned} \quad T_{\text{zero}} \leq T_{\text{total}} < T_{\text{total,inf,max}}, \quad (\text{D.33})$$

$$\begin{aligned} T_{\text{ST1}}^* &\in \left[\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1}, \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \right], \\ T_{\text{ST2}}^* &\in \left[\frac{T_{\text{total}} - \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1}{i_2(q)}, \phi_{2,\omega_2,q}^{-1}(T_{2,\text{inf,max}}) \right], \end{aligned} \quad T_{\text{zero}} \leq T_{\text{total}} < T_{\text{total,inf,max}}. \quad (\text{D.34})$$

Proof of Lemma 3:

Step I: Assume $u_1^* > T_{1,\text{inf,max}}$. By doing so, (4.16) determines $u_2^* < 0$ (torque of the EM2). Choose a small increment $\delta u_1 > 0$ in such a way that $u_1^* - \delta u_1 > T_{1,\text{inf,max}}$. The equality constraint (4.16) determines the new torque of EM2 corresponding to $u_1^* - \delta u_1$ and is denoted as $u_2^* + \delta u_2$, $\delta u_2 > 0$, which satisfies $u_2^* + \delta u_2 < 0$. By optimality,

$$l_{x,q}((u_1 - \delta u_1, u_2 + \delta u_2)^T) \geq l_{x,q}((u_1^*, u_2^*)^T). \quad (\text{D.35})$$

Consider the deviation between the left and right sides of (D.35) and follow the steps in the proof of Lemma 2, this gives a contradiction. Therefore, $u_1^* \leq T_{1,\text{inf,max}}$. Because of the equality constraint

$$(4.16), u_2^* \geq \phi_{2,\omega_2,q}^{-1} \left(\frac{T_{\text{total}} - \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1}{i_2(q)} \right).$$

Step II: Assume $u_2^* > T_{2,\text{inf,max}}$. By doing so, (4.16) determines $u_1^* < 0$. Carry on the rest of Step I, a contradiction can be found. Therefore, $u_2^* \leq T_{2,\text{inf,max}}$, which gives the lower bound for $u_1^* \geq \phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1} \right)$.

Step III: Show the range is not empty. The deviation between the upper and lower bounds is

$$T_{1,\text{inf,max}} - \phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1} \right). \quad (\text{D.36})$$

Since ϕ is monotonic, (D.36) is bigger than zero if

$$\phi_{1,\omega_1}(T_{1,\text{inf,max}}) > \frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1}, \quad (\text{D.37})$$

which is true, since

$$\frac{\phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1 + \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q) - T_{\text{total}}}{i_1} > \frac{T_{\text{total,inf,max}} - T_{\text{total}}}{i_1} > 0. \quad (\text{D.38})$$

Therefore, the range $[\phi_{1,\omega_1}^{-1} \left(\frac{T_{\text{total}} - \phi_{2,\omega_2,q}(T_{2,\text{inf,max}}) \cdot i_2(q)}{i_1} \right), T_{1,\text{inf,max}}]$ is not empty. Same can be done for the range $[\phi_{2,\omega_2,q}^{-1} \left(\frac{T_{\text{total}} - \phi_{1,\omega_1}(T_{1,\text{inf,max}}) \cdot i_1}{i_2(q)} \right), T_{2,\text{inf,max}}]$.

■

Appendix E The Bearings and Gears in Speed4E Powertrain

Tooth flank pitting

The calculated S-N curve is defined by two points: 1) the boundary stress σ_{HG} , i.e. the stress under which the fatigue strength zone starts, and the boundary cycle number N_{HG} ; 2) the fatigue limit stress σ_{HD} and the fatigue limit cycle number N_{HD} . σ_{HG} is determined by

$$\sigma_{HG} = \sigma_{HD} \cdot Z_{NT} \cdot Z_L \cdot Z_R \cdot Z_V \cdot Z_W \cdot Z_X, \quad (9.39)$$

where the parameters are defined in the following table.

Table E. 1 Parameters to calculate σ_{HG}

| Parameter | Meaning | Source |
|---------------|--------------------------------|------------------------------|
| N_{HG} | boundary cycle number | DIN 3990—Part 2, Pages 12-13 |
| σ_{HD} | fatigue limit stress | DIN 3990—Part 5, Pages 3 |
| N_{HD} | fatigue limit cycle number | DIN 3990—Part 5, Pages 3 |
| Z_{NT} | service life factor | DIN 3990—Part 2, Pages 11–12 |
| Z_L | lubricant factor | DIN 3990—Part 2, Page 13 |
| Z_R | roughness factor | DIN 3990—Part 2, Page 15 |
| Z_V | velocity factor | DIN 3990—Part 2, Page 14 |
| Z_W | material mating factor | DIN 3990—Part 2, Page 16 |
| Z_X | size factor for surface stress | DIN 3990—Part 2, Page 1 |

Table E. 2: Bearings used in Speed4E powertrain

| Bearing type | Bearing label |
|------------------------------|--|
| Hybrid bearing ⁶⁴ | B1.1.1, B1.1.2, B2.1.1, B2.1.2, B2.1.3, B2.1.4 |
| Needle bearing | B1.2.1, B1.2.2, B1.2.3 |

⁶⁴ Ceramic rolling elements and steel rings. They are essential for applications with high rotational frequency.

| | |
|---------------------|--|
| Cylinder bearing | B2.2.1, B2.3.1 |
| Deep groove bearing | B1.3.1, B1.3.2, B2.2.2, B2.3.2, B3.1.1, B3.1.2 |

Table E. 3: Parameters of the gears in Speed4E powertrain.

| | G1.S1.SG | G1.S1.PG | G1.S1.RG | G1.S2.P | G1.S2.W |
|-----------------------------|----------|----------|----------|----------|----------|
| Number of teeth | 25 | 62 | 149 | 24 | 91 |
| Gear width in mm | 17 | 17 | 17 | 38 | 38 |
| Helix angle in ° | 25.65 | 25.65 | 25.65 | 23 | 23 |
| Pressure angle in ° | 17 | 17 | 17 | 20 | 20 |
| Base diameter in mm | 27.577 | 68.39 | 164.357 | 53.341 | 202.253 |
| Pitch diameter in mm | 32.1 | 74.07 | 172.2 | 63.4 | 224.182 |
| | G2.S1.P | G2.S1.W | G2.S2.1P | G2.S2.1W | G2.S2.2P |
| Number of teeth | 19 | 76 | 24 | 74 | 34 |
| Gear width in mm | 22 | 22 | 28 | 28 | 27 |
| Helix angle in ° | 30 | 30 | 30 | 30 | 30 |
| Pressure angle in ° | 17.5 | 17.5 | 17.5 | 17.5 | 17.5 |
| Base diameter in mm | 32.985 | 141.939 | 58.383 | 164.599 | 136.942 |
| Pitch diameter in mm | 40.75 | 143.1 | 62.8 | 177.5 | 150.25 |
| | G2.S2.2W | G2.S3.P | G2.S3.W | | |
| Number of teeth | 59 | 31 | 91 | | |
| Gear width in mm | 27 | 38 | 38 | | |
| Helix angle in ° | 30 | 23 | 23 | | |
| Pressure angle in ° | 17.5 | 20 | 20 | | |
| Base diameter in mm | 78.916 | 70.092 | 202.253 | | |
| Pitch diameter in mm | 90.5 | 79.8 | 224.182 | | |

Table E. 4: Parameters of the bearings in Speed4E powertrain

| | B1.1.1 | B1.1.2 | B1.2.x | B1.3.1 | B1.3.2 | B2.1.1 | B2.1.2 |
|--------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Basic dynamic load rating in N | 13700 | 14400 | 22000 | 18500 | 35500 | 13700 | 14400 |
| Basic static load rating in N | 5600 | 6400 | 32000 | 16100 | 19100 | 5600 | 6400 |
| Fatigue limit load in N | 455 | 520 | 5700 | 890 | 1290 | 455 | 520 |
| Service life exponent | 3 | 3 | 10/3 | 3 | 3 | 3 | 3 |
| | B2.1.3 | B2.1.4 | B2.2.1 | B2.2.2 | B2.3.1 | B2.3.2 | |
| Basic dynamic load rating in N | 13700 | 14400 | 26500 | 35500 | 86000 | 35500 | |
| Basic static load rating in N | 5600 | 6400 | 11500 | 19100 | 75000 | 19100 | |
| Fatigue limit load in N | 455 | 520 | 790 | 1290 | 13200 | 1290 | |
| Service life exponent | 3 | 3 | 3 | 3 | 10/3 | 3 | |
| | B3.1.1 | B3.1.2 | | | | | |
| Basic dynamic load rating in N | 65000 | 65000 | | | | | |
| Basic static load rating in N | 21000 | 21000 | | | | | |
| Fatigue limit load in N | 1070 | 1070 | | | | | |
| Service life exponent | 3 | 3 | | | | | |

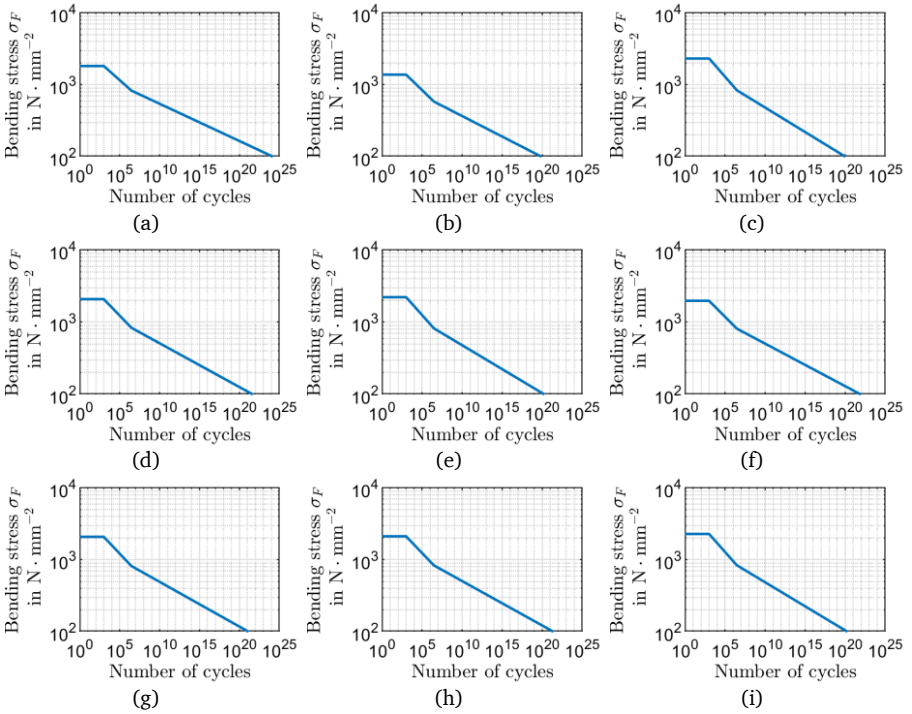


Figure E. 1 Modified S-N Curves of tooth root breakage. (a) G1.S1.SG, (b) G1.S1.PG, (c) G1.S1.RG, (d) G1.S2.P, (e) G2.S1.P, (f) G2.S2.1P, (g) G2.S2.2P, (h) G2.S3.P and (i) G.FD.W.

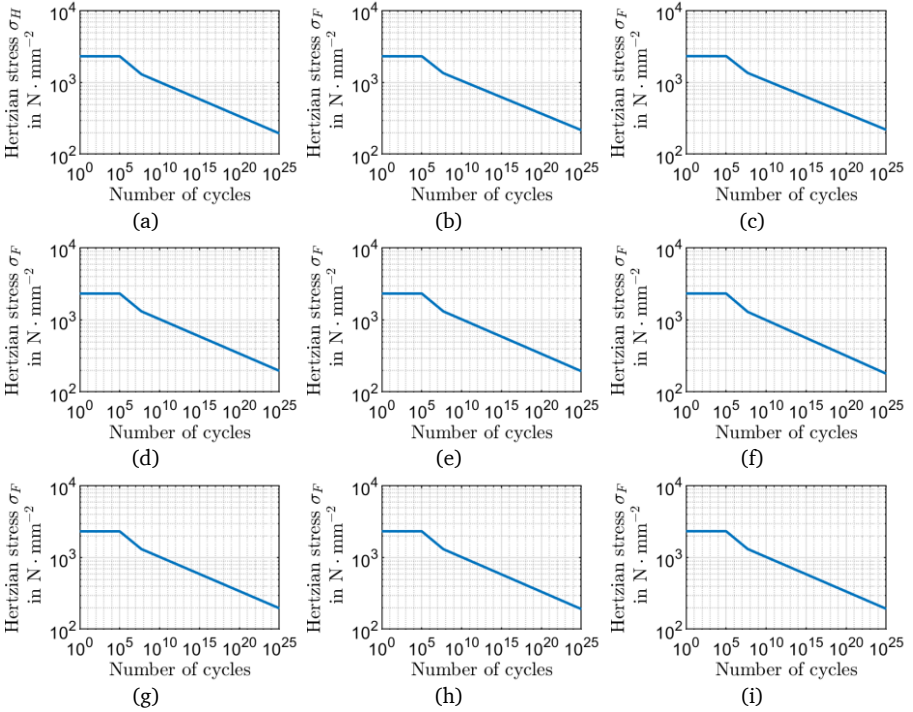


Figure E. 2 Modified S-N Curves of tooth flank pitting. (a) G1.S1.SG, (b) G1.S1.PG, (c) G1.S1.RG, (d) G1.S2.P, (e) G2.S1.P, (f) G2.S2.1P, (g) G2.S2.2P, (h) G2.S3.P and (i) G.FD.W.

Appendix F Multi-Criteria HOCP with Option 3

The continuous states for the HOCP with Option 3 in (8.24) are $\mathbf{x} := (\omega_1, \omega_2, D_{B,SPT1}, D_{B,SPT2})^T$. Note that the definition of \mathbf{x} is only valid in this appendix. The rest of the variables follow the original HOCP in Section 5.2.2. The multi-criteria cost functional with Option 3 is formulated as

$$J_{\text{multi}} = m_{\text{multi}}(\mathbf{x}(t_f)) + \int_{t_0}^{t_f} l(\mathbf{x}(t), \mathbf{u}(t)) dt, \quad (\text{F.1})$$

$$m_{\text{multi}}(\mathbf{x}(t_f)) = \beta \left\| \left(\omega_1(t_f), \omega_2(t_f) \right)^T - \left(\omega_{1,t_f}, \omega_{2,t_f} \right)^T \right\|^2 + \gamma \cdot \max\{D_{B,SPT1}, D_{B,SPT2}\}. \quad (\text{F.2})$$

The indexed Hamiltonian of the HOCP with J_{multi} is

$$\begin{aligned} \mathcal{H}_{\text{multi},q}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_{\text{multi},q}) &= \boldsymbol{\lambda}_{\text{multi},q}^T \cdot \dot{\mathbf{x}} + l((\omega_1, \omega_2)^T, \mathbf{u}) \\ &= \lambda_{\text{multi},1,q} \cdot \dot{\omega}_1 + \lambda_{\text{multi},2,q} \cdot \dot{\omega}_2 + \lambda_{\text{multi},3,q} \cdot d_{B,SPT1} \\ &\quad + \lambda_{\text{multi},4,q} \cdot d_{B,SPT2} + l((\omega_1, \omega_2)^T, \mathbf{u}) \\ &= \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q) + \lambda_{\text{multi},3,q} \cdot d_{B,SPT1}(\omega_1, u_1) + \lambda_{\text{multi},4,q} \\ &\quad \cdot d_{B,SPT2}(\omega_2, u_2), \end{aligned} \quad (\text{F.3})$$

where $\mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q)$ and $\boldsymbol{\lambda}_q = (\lambda_{1,q}, \lambda_{2,q})^T$ are the indexed Hamiltonian and the costates of the original HOCP described in Section 5.2.2. See Section 8.2.1.1 for the definition of the function d . $\boldsymbol{\lambda}_{\text{multi},q}$ denotes the indexed costates of the HOCP with J_{multi} , whose dynamics are

$$\dot{\boldsymbol{\lambda}}_{\text{multi},q}(t) = (\dot{\lambda}_{\text{multi},1,q}, \dot{\lambda}_{\text{multi},2,q}, \dot{\lambda}_{\text{multi},3,q}, \dot{\lambda}_{\text{multi},4,q})^T \quad (\text{F.4})$$

$$= - \frac{\partial \mathcal{H}_{\text{multi},q}}{\partial \mathbf{x}_q}(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}_{\text{multi},q}), \quad (\text{F.5})$$

$$\dot{\lambda}_{\text{multi},1,q} = \dot{\lambda}_{1,q}, \quad (\text{F.6})$$

$$\dot{\lambda}_{\text{multi},2,q} = \dot{\lambda}_{2,q}, \quad (\text{F.7})$$

$$\dot{\lambda}_{\text{multi},3,q} = 0, \quad (\text{F.8})$$

$$\dot{\lambda}_{\text{multi},4,q} = 0. \quad (\text{F.9})$$

$\lambda_{\text{multi},3,q}(\cdot)$ and $\lambda_{\text{multi},4,q}(\cdot)$ are two horizontal lines. Their value can be determined by the transversality condition, which is

$$\boldsymbol{\lambda}_{\text{multi},q}(t_f) = \nabla_{\mathbf{x}} m_{\text{multi}}(\mathbf{x}(t_f)). \quad (\text{F.10})$$

The first two entries of $\boldsymbol{\lambda}_{\text{multi},q}(t_f)$ are the same as discussed in Section 4.5. The third and fourth

ones are

$$\begin{aligned}\lambda_{\text{multi},3,q}(t_f) &= \gamma \cdot \frac{\partial}{\partial D_{\text{B,SPT1}}} \max\{D_{\text{B,SPT1}}(t_f), D_{\text{B,SPT2}}(t_f)\} \\ &= \begin{cases} 0, & D_{\text{B,SPT1}}(t_f) < D_{\text{B,SPT2}}(t_f), \\ \gamma, & D_{\text{B,SPT1}}(t_f) \geq D_{\text{B,SPT2}}(t_f), \end{cases} \end{aligned} \quad (\text{F.11})$$

$$\begin{aligned}\lambda_{\text{multi},4,q}(t_f) &= \gamma \cdot \frac{\partial}{\partial D_{\text{B,SPT2}}} \max\{D_{\text{B,SPT1}}(t_f), D_{\text{B,SPT2}}(t_f)\} \\ &= \begin{cases} \gamma, & D_{\text{B,SPT1}}(t_f) \leq D_{\text{B,SPT2}}(t_f), \\ 0, & D_{\text{B,SPT1}}(t_f) > D_{\text{B,SPT2}}(t_f), \end{cases} \end{aligned} \quad (\text{F.12})$$

which indicates that

$$\lambda_{\text{multi},3,q}(\cdot) = \begin{cases} 0, & D_{\text{B,SPT1}}(t_f) < D_{\text{B,SPT2}}(t_f), \\ \gamma, & D_{\text{B,SPT1}}(t_f) \geq D_{\text{B,SPT2}}(t_f), \end{cases} \quad (\text{F.13})$$

$$\lambda_{\text{multi},4,q}(\cdot) = \begin{cases} \gamma, & D_{\text{B,SPT1}}(t_f) \leq D_{\text{B,SPT2}}(t_f), \\ 0, & D_{\text{B,SPT1}}(t_f) > D_{\text{B,SPT2}}(t_f). \end{cases} \quad (\text{F.14})$$

Importantly, $D_{\text{B,SPT1}}$ and $D_{\text{B,SPT2}}$ at t_f are unknown.

Replace $\lambda_{\text{multi},3,q}$ and $\lambda_{\text{multi},4,q}$ in (F.3) to obtain

$$\begin{aligned}\mathcal{H}_{\text{multi},q}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_{\text{multi},q}) &= \begin{cases} \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q) + \gamma \cdot d_{\text{B,SPT2}}(\omega_2, u_2), & D_{\text{B,SPT1}}(t_f) < D_{\text{B,SPT2}}(t_f), \\ \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q) + \gamma \cdot d_{\text{B,SPT1}}(\omega_1, u_1) + \gamma \cdot d_{\text{B,SPT2}}(\omega_1, u_1), & D_{\text{B,SPT1}}(t_f) = D_{\text{B,SPT2}}(t_f), \\ \mathcal{H}_q(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}_q) + \gamma \cdot d_{\text{B,SPT1}}(\omega_1, u_1), & D_{\text{B,SPT1}}(t_f) > D_{\text{B,SPT2}}(t_f). \end{cases} \end{aligned} \quad (\text{F.15})$$

The condition depends on the states at t_f , which are unknown in $t \in [t_0, t_f)$. There is no evidence to show that $D_{\text{B,SPT1}}(t_f) = D_{\text{B,SPT2}}(t_f)$ and, therefore, one of the condition 1 and 3 has to be randomly chosen without any confidence to initialize the costates in (F.13) and (F.14) and determine the formulation of the Hamiltonian in (F.15). The minimization of the instantaneous Hamiltonian only considers either $d_{\text{B,SPT1}}$ or $d_{\text{B,SPT2}}$, which causes the relation between $D_{\text{B,SPT1}}(t_f)$ and $D_{\text{B,SPT2}}(t_f)$ inevitably contrary to the condition considered for the Hamiltonian. A convergence is difficult to reach, if ever possible. One possible work-around is to assume that $D_{\text{B,SPT1}}(t_f) = D_{\text{B,SPT2}}(t_f)$ for all cases, which makes (F.15) equivalent to the multi-criteria cost functional with Option 2 in Section 8.2.1.1.

Appendix G Fatigue Coefficient

Table G. 1 Values of the fatigue coefficients in MCOS with the fatigue cost functional for bearings.

| | | | | | | | | | |
|------------------------------|-------------------------------|------------------------------|------------------------------|------------------------------|-------------------------------|------------------------------|-------------------------------|------------------------------|-------------------------------|
| γ_1 10^1 | γ_2 10^2 | γ_3 10^3 | γ_4 10^4 | γ_5 10^5 | γ_6 $10^{5.5}$ | γ_7 10^6 | γ_8 $10^{6.5}$ | γ_9 10^7 | γ_{10} $10^{7.5}$ |
| γ_{11} 10^8 | γ_{12} $10^{8.5}$ | γ_{13} 10^9 | γ_{14} $10^{9.5}$ | γ_{15} $10^{9.6}$ | γ_{16} $10^{9.7}$ | γ_{17} $10^{9.8}$ | γ_{18} $10^{9.9}$ | γ_{19} 10^{10} | γ_{20} $10^{10.1}$ |
| γ_{21} $10^{10.2}$ | γ_{22} $10^{10.3}$ | γ_{23} $10^{10.4}$ | γ_{24} $10^{10.5}$ | γ_{25} $10^{10.6}$ | γ_{26} $10^{10.7}$ | γ_{27} $10^{10.8}$ | γ_{28} $10^{10.9}$ | γ_{29} 10^{11} | γ_{30} $10^{11.1}$ |
| γ_{31} $10^{11.2}$ | γ_{32} $10^{11.3}$ | γ_{33} $10^{11.4}$ | γ_{34} $10^{11.5}$ | γ_{35} $10^{11.6}$ | γ_{36} $10^{11.65}$ | γ_{37} $10^{11.7}$ | γ_{38} $10^{11.75}$ | γ_{39} $10^{11.8}$ | γ_{40} $10^{11.85}$ |
| γ_{41} $10^{11.9}$ | γ_{42} $10^{11.95}$ | γ_{43} 10^{12} | γ_{44} 10^{13} | γ_{45} $10^{13.5}$ | γ_{46} 10^{14} | γ_{47} $10^{14.5}$ | γ_{48} 10^{15} | γ_{49} $10^{15.5}$ | γ_{50} 10^{16} |
| γ_{51} $10^{16.5}$ | γ_{52} 10^{17} | γ_{53} 10^{18} | γ_{54} 10^{19} | γ_{55} 10^{20} | | | | | |

Appendix H Accumulated Fatigue and Service Life

Supplementary Results for the MCOS with the Fatigue Cost Functional for Bearings

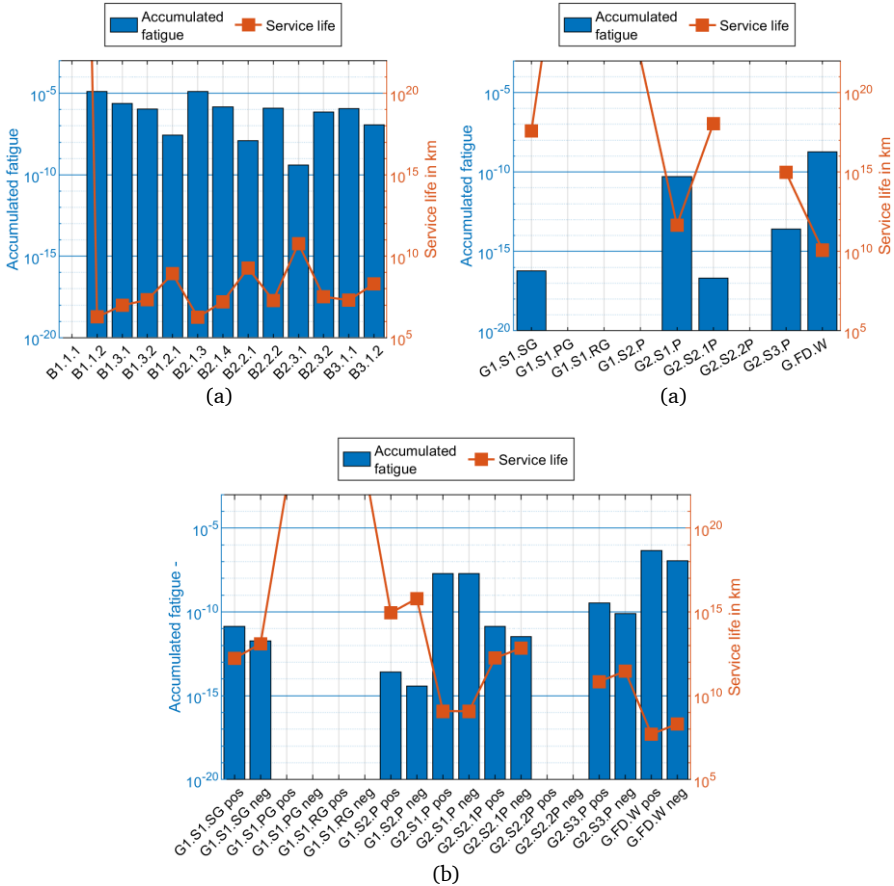


Figure H. 1: Accumulated fatigue after a WLTC and corresponding service lives of the parts in Speed4E powertrain, controlled by the MCOS with γ_{31} . (a): Bearing fatigue. (b): Tooth root breakage. (c) Tooth flank pitting.

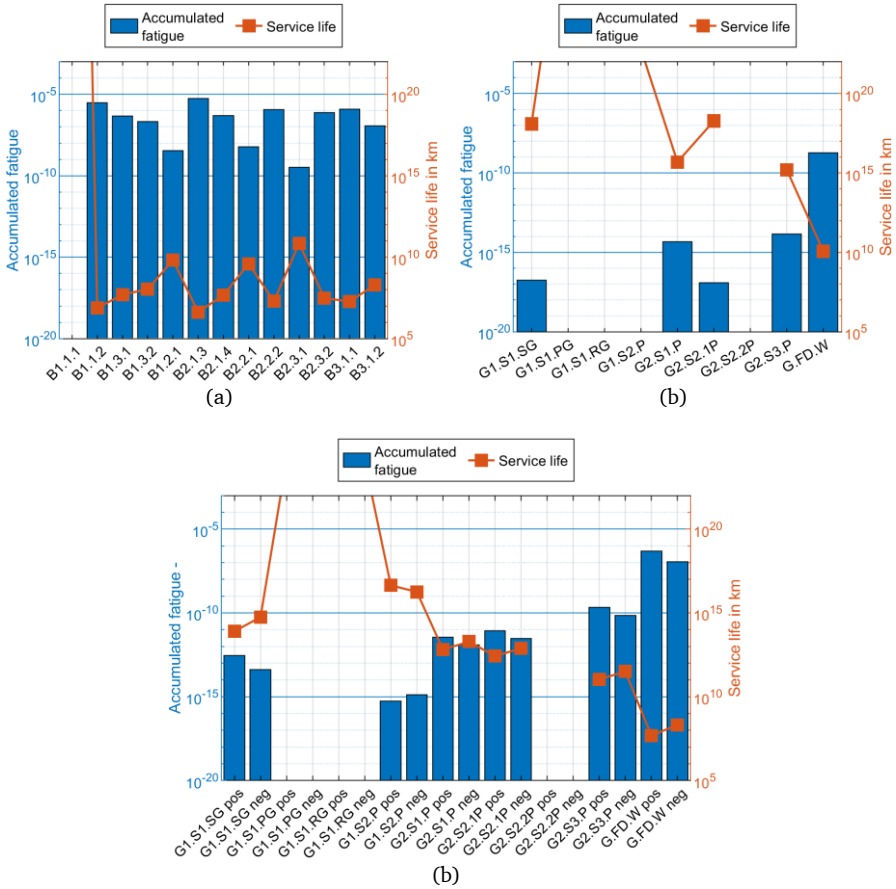


Figure H. 2: Accumulated fatigue after a WLTC and corresponding service lives of the parts in Speed4E powertrain , controlled by the MCOS with γ_{55} . (a): Bearing fatigue. (b): Tooth root breakage. (c) Tooth flank pitting.

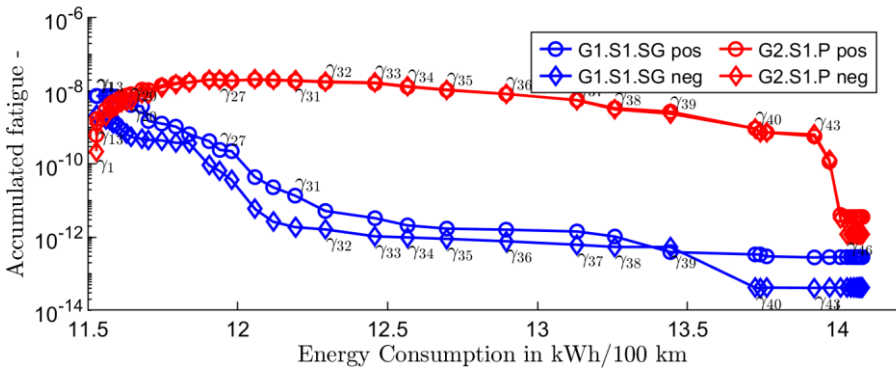


Figure H. 3: The simulation results of the MCOS with the indexed γ : Gear accumulated fatigue vs. Energy consumption.

Results for the MCOS with the Fatigue Cost Functional for Gears

The MCOS with the fatigue cost functional for bearings has been discussed in length in Sections 8.3 and 8.4. The same analysing approach can be applied to its counterpart for gears. This part does not reiterate the similar outcomes but points out the difference.

Figure H. 4 presents the values of the sampled γ for the evaluation on a logarithmic scale with the index of γ as the x-axis. The indexed γ is used as the fatigue coefficient in the MCOS with the same index. See 8.4.1 for the reason of different sampling density.

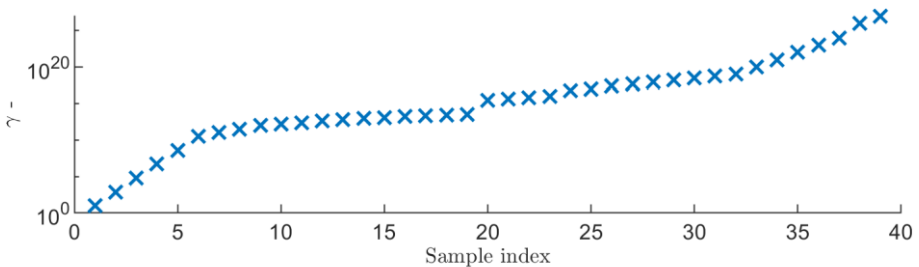


Figure H. 4: The values of γ sampled for the simulation study

Figure H. 5 summarizes the vehicle energy consumption and the gear service lives of a set of WLTC driving cycle simulations. presents the service life of G1.S1.SG, i.e. $SL_{G1.S1.SG}$ (the blue curves), and that of G2.S2.P, i.e. $SL_{G2.S1.P}$ (the red curves), on a logarithmic scale, where the circles represent the tooth flanks “pos” and the diamonds represent the tooth flanks “neg”. As γ increases

starting from γ_1 , $SL_{G2.S1.P}$ starts to decrease and reach it minimal at γ_{18} , after which it rises gradually. At the other end of the index γ , $SL_{G2.S1.P}$ rises drastically, for the same reason discussed in Section 8.4.3. From one end to the other end of γ , the gear service life considering both gears and both “pos” and “neg” extends by more than two orders of magnitude (from 3.18×10^9 km to 1.16×10^{12} km). Similar results have been achieved by the MCOS with the fatigue cost functional for bearings presented in Section 8.4.3. At the same time, 20.67 % more energy is consumed. It has extended the powertrain service life to 3.27×10^6 km, which is limited by the shortest bearing service life presented in Figure H. 6.

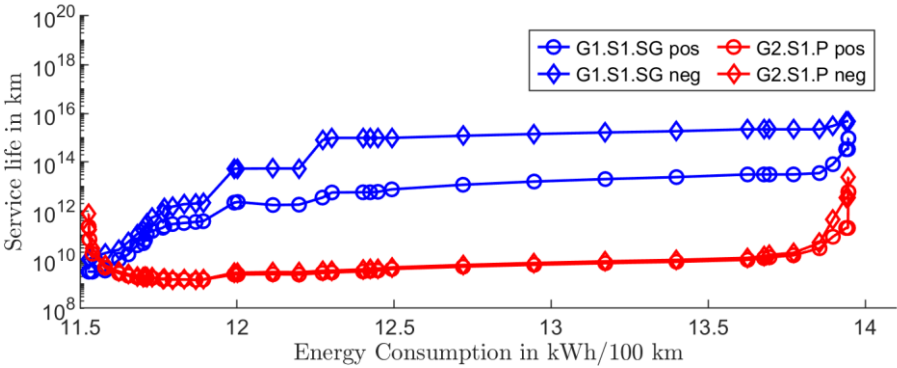


Figure H. 5: The simulation results of the MCOS with the indexed γ : Gear service life vs. Energy consumption.

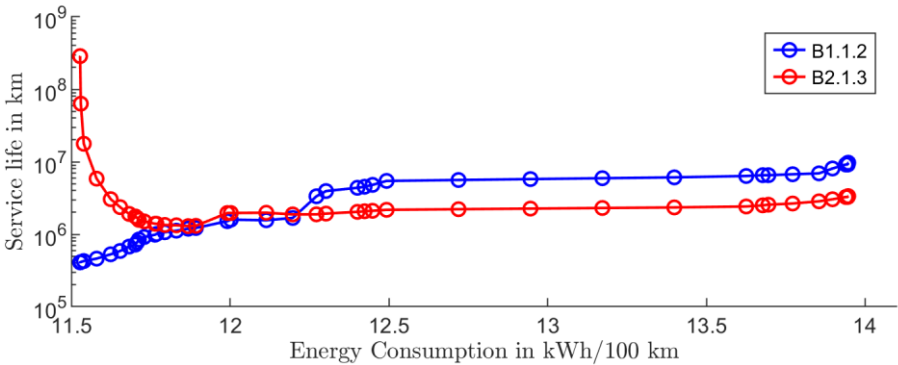


Figure H. 6: The simulation results of the MCOS with the indexed γ : Bearing service life vs. Energy consumption.

References

- [1] European Commission, *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: The European Green Deal*. Brussels, 2019. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1576150542719&uri=COM%3A2019%3A640%3AFIN>
- [2] United Nations Framework Convention on Climate Change. "A Beginner's Guide to Climate Neutrality." [Online]. Available: <https://unfccc.int/blog/a-beginner-s-guide-to-climate-neutrality#:~:text=Climate%20neutrality%20refers%20to%20the,our%20emissions%20through%20climate%20action.>
- [3] International Energy Agency, *Tracking Clean Energy Progress 2023*. Paris, 2023. [Online]. Available: <https://www.iea.org/reports/tracking-clean-energy-progress-2023>
- [4] Energy Technology Policy Division of the Directorate of Sustainability, Technology and Outlooks of the International Energy Agency, "Global EV Outlook 2023: Catching up with climate ambitions," International Energy Agency, Rep. 0-142, 2023. [Online]. Available: <https://iea.blob.core.windows.net/assets/dacf14d2-eabc-498a-8263-9f97fd5dc327/GEVO2023.pdf>
- [5] European Parliament and the Council of the European Union, *Regulation (EU) 2023/851 of the European Parliament and of the Council of 19 April 2023 amending Regulation (EU) 2019/631 as regards strengthening the CO2 emission performance standards for new passenger cars and new light commercial vehicles in line with the Union's increased climate ambition (Text with EEA relevance)*, 2023. [Online]. Available: <http://data.europa.eu/eli/reg/2023/851/oj>
- [6] I. Frenzel, J. Jarass, S. Trommer, and B. Lenz, "Erstnutzer von Elektrofahrzeugen in Deutschland. Nutzerprofile, Anschaffung, Fahrzeugnutzung," Institut für Verkehrsfor-schung, Berlin, Rep. 1-69, 26. May, 2015. [Online]. Available: https://elib.dlr.de/96491/1/Ergebnisbericht_E-Nutzer_2015.pdf
- [7] X. Zhou, P. Walker, N. Zhang, and B. Zhu, "Performance Improvement of a Two Speed EV through Combined Gear Ratio and Shift Schedule Optimization," *SAE 2013 World Congress & Exhibition*, 2013, doi: 10.4271/2013-01-1477.
- [8] Z. Zhang, C. Zuo, W. Hao, Y. Zuo, X. L. Zhao, and M. Zhang, "Three-speed transmission

system for purely electric vehicles," *Int. J. Automot. Technol.*, vol. 14, no. 5, pp. 773–778, 2013, doi: 10.1007/s12239-013-0085-0.

- [9] B. Gao, Q. Liang, Y. Xiang, L. Guo, and H. Chen, "Gear ratio optimization and shift control of 2-speed I-AMT in electric vehicle," *Mechanical Systems and Signal Processing*, vol. 50-51, pp. 615–631, 2015, doi: 10.1016/j.ymsp.2014.05.045.
- [10] M. R. Ahssan, M. M. Ektesabi, and S. A. Gorji, "Electric Vehicle with Multi-Speed Transmission: A Review on Performances and Complexities," *SAE Int. J. Alt. Power.*, vol. 7, no. 2, pp. 169–181, 2018, doi: 10.4271/08-07-02-0011.
- [11] C. T.P. Nguyen, B.-H. Nguyễn, J. P. F. Trovão, and M. C. Ta, "Optimal drivetrain design methodology for enhancing dynamic and energy performances of dual-motor electric vehicles," *Energy Conversion and Management*, vol. 252, p. 115054, 2022, doi: 10.1016/j.enconman.2021.115054.
- [12] Z. Wang, J. Zhou, and G. Rizzoni, "A review of architectures and control strategies of dual-motor coupling powertrain systems for battery electric vehicles," *Renewable and Sustainable Energy Reviews*, vol. 162, p. 112455, 2022, doi: 10.1016/j.rser.2022.112455.
- [13] J. Wu, J. Liang, J. Ruan, N. Zhang, and P. D. Walker, "Efficiency comparison of electric vehicles powertrains with dual motor and single motor input," *Mechanism and Machine Theory*, vol. 128, pp. 569–585, 2018, doi: 10.1016/j.mechmachtheory.2018.07.003.
- [14] P. Pisu and G. Rizzoni, "A Comparative Study Of Supervisory Control Strategies for Hybrid Electric Vehicles," *IEEE Trans. Contr. Syst. Technol.*, vol. 15, no. 3, pp. 506–518, 2007, doi: 10.1109/TCST.2007.894649.
- [15] L. Serrao *et al.*, "Open Issues in Supervisory Control of Hybrid Electric Vehicles: A Unified Approach Using Optimal Control Methods," *Oil & Gas Science and Technology*, vol. 68, no. 1, pp. 23–33, 2013, doi: 10.2516/ogst/2012080.
- [16] D. E. Kirk, *Optimal control theory: An introduction*. Mineola, N.Y.: Dover Publications, 2004.
- [17] J. Oldenburg, W. Marquardt, D. Heinz, and D. B. Leineweber, "Mixed-logic dynamic optimization applied to batch distillation process design," *AIChE J.*, vol. 49, no. 11, pp. 2900–2917, 2003, doi: 10.1002/aic.690491120.
- [18] S. Sager, *Numerical methods for mixed-integer optimal control problems* (Zugl.: Heidelberg, Univ., Diss., [2006]). Tönning, Lübeck, Marburg: Der Andere Verl., 2005.
- [19] M. Buss, M. Glocker, M. Hardt, O. von Stryk, R. Bulirsch, and G. Schmidt, "Nonlinear Hy-

brid Dynamical Systems: Modeling, Optimal Control, and Applications," in *Modelling, analysis, and design of hybrid systems* (Engineering online library 279), S. Engell, G. Frehse, and E. Schnieder, Eds., Berlin, Heidelberg: Springer, 2002, pp. 311–335.

- [20] H. J. Sussmann, "A maximum principle for hybrid optimal control problems," in *Proceedings of the 38th IEEE Conference on Decision and Control: December 7 - 10, 1999, Crowne Plaza Hotel & Resort, Phoenix, Arizona, USA, 1999*, pp. 425–430, doi: 10.1109/CDC.1999.832814.
- [21] 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, doi: 10.3390/en6042305.
- [22] G. Paganelli, S. Delprat, T. M. Guerra, J. Rimaux, and J. J. Santin, "Equivalent consumption minimization strategy for parallel hybrid powertrains," *Vehicular Technology Conference. IEEE 55th Vehicular Technology Conference. VTC Spring 2002 (Cat. No.02CH37367), Birmingham, AL, USA*, pp. 2076–2081, 2002, doi: 10.1109/VTC.2002.1002989.
- [23] L. Serrao, S. Onori, and G. Rizzoni, "ECMS as a realization of Pontryagin's minimum principle for HEV control," in *2009 American Control Conference: ACC 2009 ; St. Louis, Missouri, USA, 10 - 12 June 2009 ; [annual conference of the American Automatic Control Council, 2009*, pp. 3964–3969, doi: 10.1109/ACC.2009.5160628.
- [24] A. Brahma, Y. Guezennec, and G. Rizzoni, "Optimal energy management in series hybrid electric vehicles," in *Proceedings of the 2000 American Control Conference, ACC: June 28 - 30, 2000, Hyatt Regency Chicago, Chicago, Illinois, USA*, Jim Zhu, Ed., vol. 1, 2000, 60-64, doi: 10.1109/ACC.2000.878772.
- [25] 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. Contr. Syst. Technol.*, vol. 20, no. 6, pp. 1490–1505, 2012, doi: 10.1109/TCST.2011.2168820.
- [26] A. Pakniyat and P. E. Caines, "Hybrid optimal control of an electric vehicle with a dual-planetary transmission," *Nonlinear Analysis: Hybrid Systems*, vol. 25, pp. 263–282, 2017, doi: 10.1016/j.nahs.2016.08.004.
- [27] Peng Dong, *Optimized Shift Control in Automatic Transmissions with respect to Spontaneity, Comfort, and Shift Loads* (Zugl. : Ruhr-Universität Bochum, Diss., [2015]). Bochum: Ruhr-Universität Bochum, 2015.
- [28] A. Haj-Fraj and F. Pfeiffer, "Optimal control of gear shift operations in automatic transmissions," *Journal of the Franklin Institute*, vol. 338, 2-3, pp. 371–390, 2001, doi:

10.1016/S0016-0032(00)00091-0.

- [29] S. Kim, J. Oh, and S. Choi, "Gear shift control of a dual-clutch transmission using optimal control allocation," *Mechanism and Machine Theory*, vol. 113, pp. 109–125, 2017, doi: 10.1016/j.mechmachtheory.2017.02.013.
- [30] M. Kulkarni, T. Shim, and Y. Zhang, "Shift dynamics and control of dual-clutch transmissions," *Mechanism and Machine Theory*, vol. 42, no. 2, pp. 168–182, 2007, doi: 10.1016/j.mechmachtheory.2006.03.002.
- [31] H. Zhang, Y. Zhang, C. Hao, and Y. Lu, "Test Research on Optimization of AMT Shift Control Parameters," in *2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI): Oct. 28-30, 2022, Nanjing, China : conference digest*, 2022, pp. 1–6, doi: 10.1109/CVCI54083.2021.9661185.
- [32] Z. Zhong, G. Kong, Z. Yu, X. Xin, and X. Chen, "Shifting control of an automated mechanical transmission without using the clutch," *Int.J Automot. Technol.*, vol. 13, no. 3, pp. 487–496, 2012, doi: 10.1007/s12239-012-0046-z.
- [33] Y. Zou, H. Shi-jie, L. Dong-ge, G. Wei, and X.-s. Hu, "Optimal Energy Control Strategy Design for a Hybrid Electric Vehicle," *Discrete Dynamics in Nature and Society*, vol. 2013, pp. 1–8, 2013, doi: 10.1155/2013/132064.
- [34] C.-C. Lin, H. Peng, J. W. Grizzle, and J.-M. Kang, "Power management strategy for a parallel hybrid electric truck," *IEEE Trans. Contr. Syst. Technol.*, vol. 11, no. 6, pp. 839–849, 2003, doi: 10.1109/TCST.2003.815606.
- [35] 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," *Applied Energy*, vol. 236, pp. 893–905, 2019, doi: 10.1016/j.apenergy.2018.12.032.
- [36] K. Trieb, H. Naunheimer, B. Bertsche, J. Ryborz, W. Novak, and P. Fietkau, *Revisionsendoprothetik des Kniegelenks // Automotive Transmissions: Fundamentals, Selection, Design and Application*, 2nd ed. (SpringerLink Bücher). Berlin, Heidelberg: Springer-Verlag GmbH Berlin Heidelberg; Springer Berlin Heidelberg, 2011.
- [37] Y. Tao *et al.*, "Model Predictive Energy Control Including Mechanical Fatigue Life of a Two-Motor Multi-Speed Electric Vehicle," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 14109–14116, 2020, doi: 10.1016/j.ifacol.2020.12.1009.
- [38] A. Pakniyat and P. E. Caines, "On the Hybrid Minimum Principle: The Hamiltonian and Adjoint Boundary Conditions," *IEEE Trans. Automat. Contr.*, vol. 66, no. 3, pp. 1246–1253, 2021, doi: 10.1109/TAC.2020.2992450.

- [39] W.P.M.H. Heemels, P. Bernard, K.J.A. Scheres, R. Postoyan, and R. G. Sanfelice, "Hybrid Systems with Continuous-time Inputs: Subtleties in Solution Concepts and Existence Results," *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 5368–5373, 2021, doi: 10.1109/CDC45484.2021.9683389.
- [40] Ali Pakniyat, "Optimal Control of Deterministic and Stochastic Hybrid Systems: Theory and Applications," Dissertation, Department of Electrical & Computer Engineering, McGill University, Montreal, Canada, 2016.
- [41] B. Passenberg, "Theory and algorithms for indirect methods in optimal control of hybrid systems," PhD thesis, Lehrstuhl für Steuerungs- und Regelungstechnik, Technischen Universität München, Munich, Germany, 2011.
- [42] M. S. Branicky, V. S. Borkar, and S. K. Mitter, "A unified framework for hybrid control: model and optimal control theory," *IEEE Trans. Automat. Contr.*, vol. 43, no. 1, pp. 31–45, 1998, doi: 10.1109/9.654885.
- [43] X. Xu and P. J. Antsaklis, "Results and Perspectives on Computational Methods for Optimal Control of Switched Systems," vol. 2623, pp. 540–555, 2003, doi: 10.1007/3-540-36580-X_39.
- [44] H. L. Royden and P. Fitzpatrick, *Real Analysis*. Boston: Prentice Hall, 2010.
- [45] A. D. Ioffe and V. M. Tichomirov, *Theory of extremal problems* (Studies in mathematics and its applications 6). Amsterdam: North-Holland Publ.Co, 1979.
- [46] T. Tao, *Analysis* (Texts and readings in mathematics 37). New Delhi, India: Hindustan Book Agency, 2022.
- [47] W.P.M.H. Heemels, M. K. Çamhbel, A. J. van der Schaft, and J. M. Schumacher, "On the existence and uniqueness of solution trajectories to hybrid dynamical systems," *Nonlinear and Hybrid Control in Automotive Applications*, pp. 391–422, 2003.
- [48] J. C. Butcher, "A history of Runge-Kutta methods," *Applied Numerical Mathematics*, vol. 20, no. 3, pp. 247–260, 1996, doi: 10.1016/0168-9274(95)00108-5.
- [49] W. W. Hager, "Runge-Kutta methods in optimal control and the transformed adjoint system," *Numerische Mathematik*, vol. 87, no. 2, pp. 247–282, 2000, doi: 10.1007/s002110000178.
- [50] J. F. Bonnans and J. Laurent-Varin, "Computation of order conditions for symplectic partitioned Runge-Kutta schemes with application to optimal control," *Numerische Mathematik*, vol. 103, no. 1, pp. 1–10, 2006, doi: 10.1007/s00211-005-0661-y.

- [51] A. Walther, "Automatic differentiation of explicit Runge-Kutta methods for optimal control," *Comput Optim Applic*, vol. 36, no. 1, pp. 83–108, 2007, doi: 10.1007/s10589-006-0397-3.
- [52] Dirk Schröder, "Peer Methods in Optimal Control," Ph.D. Dissertation, Fachbereich Mathematik, Technischen Universität Darmstadt, Darmstadt, 2016. [Online]. Available: https://tuprints.ulb.tu-darmstadt.de/5440/1/schroeder_dissertation_final.pdf
- [53] E. Hairer and G. Wanner, *Stiff and differential-algebraic problems*, 2nd ed. (Springer series in computational mathematics 14). Berlin: Springer, 1996.
- [54] R. Bellman, *Dynamic programming*, 1st ed. (Princeton Landmarks in mathematics). Princeton, NJ: Princeton University Press, 2010.
- [55] D. P. Bertsekas, *Dynamic programming and optimal control* (Athena scientific optimization and computation series 1). Belmont, Massachusetts: Athena Scientific, 2017.
- [56] M. Rungger, "On the Numerical Solution of Nonlinear and Hybrid Optimal Control Problems," Ph.D. dissertation, Kassel University, Kassel, Germany, 2012. [Online]. Available: <https://www.uni-kassel.de/ub/?id=39129&h=9783862193660>
- [57] S. Hedlund and A. Rantzer, "Optimal control of hybrid systems," in *Proceedings of the 38th IEEE Conference on Decision and Control: December 7 - 10, 1999, Crowne Plaza Hotel & Resort, Phoenix, Arizona, USA*, 1999, pp. 3972–3977, doi: 10.1109/CDC.1999.827981.
- [58] X. Xu and P. J. Antsaklis, "A dynamic programming approach for optimal control of switched systems," *Proceedings of the 39th IEEE conference on decision and control*, no. 2, pp. 1822–1827, 2000, doi: 10.1109/CDC.2000.912127.
- [59] O. Sundström, D. Ambühl, and L. Guzzella, "On Implementation of Dynamic Programming for Optimal Control Problems with Final State Constraints," *Oil & Gas Science and Technology*, vol. 65, no. 1, pp. 91–102, 2010, doi: 10.2516/ogst/2009020.
- [60] J. Bernoulli, "A new problem to whose solution mathematicians are invited (Problema novum ad cuius solutionem Mathematici invitantur)," *Acta Eruditorum*, no. 18, p. 269, 1696. [Online]. Available: <https://books.google.com/books?id=4q1RAAAcAAJ&pg=PA269>
- [61] H. J. Sussmann and J. C. Willems, "300 years of optimal control: from the brachystochrone to the maximum principle," *IEEE Control Syst.*, vol. 17, no. 3, pp. 32–44, 1997, doi: 10.1109/37.588098.
- [62] L. S. Pontryagin, *The Mathematical Theory of Optimal Processes* (4). New York: Gordon and Breach, 1986.

- [63] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, New York, NY, Port Melbourne: Cambridge University Press, 2004.
- [64] P. Riedinger, C. lung, and F. Kratz, "An Optimal Control Approach for Hybrid Systems," *European Journal of Control*, vol. 9, no. 5, pp. 449–458, 2003, doi: 10.3166/ejc.9.449-458.
- [65] B. Passenberg, M. Leibold, O. Stursberg, and M. Buss, "The minimum principle for time-varying hybrid systems with state switching and jumps," in *2011 50th IEEE Conference on Decision and Control and European Control Conference: (CDC-ECC) ; 12 - 15 Dec. 2011, Orlando, FL, USA*, 2011, pp. 6723–6729, doi: 10.1109/CDC.2011.6160587.
- [66] M. S. Shaikh, "Optimal control of hybrid systems: theory and algorithms," Department of Electrical and Computer Engineering, McGill University, Montréal, Canada, 2004.
- [67] X. Xu, "Analysis and design of switched systems," Ph.D. thesis, Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana, USA, 2001.
- [68] A. V. Dmitruk and A. M. Kaganovich, "The Hybrid Maximum Principle is a consequence of Pontryagin Maximum Principle," *Systems & Control Letters*, vol. 57, no. 11, pp. 964–970, 2008, doi: 10.1016/j.sysconle.2008.05.006.
- [69] T. J. Böhme and B. Frank, *Hybrid systems, optimal control and hybrid vehicles: Theory, methods and applications* (Advances in industrial control). Cham, Switzerland: Springer, 2017.
- [70] J. Nocedal and S. J. Wright, *Numerical optimization* (Springer series in operation research and financial engineering). New York, NY: Springer, 2006.
- [71] M. Dowell and P. Jarratt, "The "Pegasus" method for computing the root of an equation," *BIT*, vol. 12, no. 4, pp. 503–508, 1972, doi: 10.1007/BF01932959.
- [72] R. Fletcher, *Practical methods of optimization*, 2nd ed. (A Wiley-Interscience Publication). Chichester: Wiley, 1987.
- [73] T. J. Boehme, M. Schori, B. Frank, M. Schultalbers, and B. Lampe, "Solution of a hybrid optimal control problem for parallel hybrid vehicles subject to thermal constraints," *52nd IEEE Conference on Decision and Control, Firenze, Italy*, pp. 2220–2226, 2013, doi: 10.1109/CDC.2013.6760211.
- [74] B. Passenberg, M. Leibold, O. Stursberg, and M. Buss, "A Globally Convergent, Locally Optimal Min- \mathcal{H}_∞ Algorithm for Hybrid Optimal Control," *SIAM J. Control Optim.*, vol. 52, no. 1, pp. 718–746, 2014, doi: 10.1137/110857969.

- [75] J. T. Betts, *Practical methods for optimal control and estimation using nonlinear programming* (Advances in design and control). Philadelphia, Pa.: Soc. for Industrial and Applied Mathematics, 2010.
- [76] J. Kronqvist, D. E. Bernal, A. Lundell, and I. E. Grossmann, "A review and comparison of solvers for convex MINLP," *Optim Eng*, vol. 20, no. 2, pp. 397–455, 2019, doi: 10.1007/s11081-018-9411-8.
- [77] N. Robuschi, C. Zeile, S. Sager, and F. Braghin, "Multiphase mixed-integer nonlinear optimal control of hybrid electric vehicles," *Automatica*, vol. 123, p. 109325, 2021, doi: 10.1016/j.automatica.2020.109325.
- [78] C. Kirches, *Fast Numerical Methods for Mixed-Integer Nonlinear Model-Predictive Control* (SpringerLink Bücher). Wiesbaden: Vieweg+Teubner Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden, 2011.
- [79] I. E. Grossmann and Z. Kravanja, "Mixed-Integer Nonlinear Programming: A Survey of Algorithms and Applications," in *Large-Scale Optimization with Applications: Part II: Optimal Design and Control* (Springer eBook Collection Mathematics and Statistics 93), T. F. Coleman, A. R. Conn, F. N. Santosa, and L. T. Biegler, Eds., New York, NY: Springer, 1997, pp. 73–100.
- [80] J. Till, S. Engell, S. Panek, and O. Stursberg, "Applied hybrid system optimization: An empirical investigation of complexity," *Control Engineering Practice*, vol. 12, no. 10, pp. 1291–1303, 2004, doi: 10.1016/j.conengprac.2004.04.003.
- [81] S. S. Skiena, *The algorithm design manual*, 2nd ed. London: Springer, 2008.
- [82] L. Schmitt, M. Keller, T. Albin, and D. Abel, "Real-Time Nonlinear Model Predictive Control for the Energy Management of Hybrid Electric Vehicles in a Hierarchical Framework," in *2020 American Control Conference: Sheraton Denver Downtown Hotel, Denver, CO, USA*, M. Oishi, Ed., 2020, pp. 1961–1967, doi: 10.23919/ACC45564.2020.9147465.
- [83] A. Panday and H. O. Bansal, "A Review of Optimal Energy Management Strategies for Hybrid Electric Vehicle," *International Journal of Vehicular Technology*, vol. 2014, pp. 1–19, 2014, doi: 10.1155/2014/160510.
- [84] S. G. Wirasingha and A. Emadi, "Classification and review of control strategies for plug-in hybrid electric vehicles," in *2009 IEEE Vehicle Power and Propulsion Conference: VPPC 2009 ; Dearborn, Michigan, USA, 7 - 11 September 2009*, 2009, pp. 907–914, doi: 10.1109/VPPC.2009.5289751.

- [85] Y. Cao, M. Yao, and X. Sun, "An Overview of Modelling and Energy Management Strategies for Hybrid Electric Vehicles," *Applied Sciences*, vol. 13, no. 10, p. 5947, 2023, doi: 10.3390/app13105947.
- [86] L. Nicoletti, F. Ostermann, M. Heinrich, A. Stauber, X. Lin, and M. Lienkamp, "Topology analysis of electric vehicles, with a focus on the traction battery," *Forsch Ingenieurwes*, vol. 85, no. 2, pp. 457–467, 2021, doi: 10.1007/s10010-020-00422-1.
- [87] J. Ruan, P. Walker, and N. Zhang, "A comparative study energy consumption and costs of battery electric vehicle transmissions," *Applied Energy*, vol. 165, pp. 119–134, 2016, doi: 10.1016/j.apenergy.2015.12.081.
- [88] W. Enang and C. Bannister, "Modelling and control of hybrid electric vehicles (A comprehensive review)," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 1210–1239, 2017, doi: 10.1016/j.rser.2017.01.075.
- [89] A. L. Dicks, "PEM Fuel Cells," in *Comprehensive renewable energy*, A. Sayigh, Ed. Amsterdam: Elsevier, 2012, pp. 203–245.
- [90] Y. Yang, X. Hu, H. Pei, and Z. Peng, "Comparison of power-split and parallel hybrid powertrain architectures with a single electric machine: Dynamic programming approach," *Applied Energy*, vol. 168, pp. 683–690, 2016, doi: 10.1016/j.apenergy.2016.02.023.
- [91] S. Delprat, T. M. Guerra, G. Paganelli, J. Lauber, and M. Delhom, "Control strategy optimization for an hybrid parallel powertrain," in *Proceedings of the 2001 American Control Conference, ACC: June 25 - 27, 2001, Crystal Gateway Marriot, Arlington, VA, USA*, 2001, 1315–1320 vol.2, doi: 10.1109/ACC.2001.945905.
- [92] W. Wang, R. Song, M. Guo, and S. Liu, "Analysis on compound-split configuration of power-split hybrid electric vehicle," *Mechanism and Machine Theory*, vol. 78, pp. 272–288, 2014, doi: 10.1016/j.mechmachtheory.2014.03.019.
- [93] Y. Gao and M. Ehsani, "A torque and speed coupling hybrid drivetrain-architecture, control, and simulation," *IEEE Trans. Power Electron.*, vol. 21, no. 3, pp. 741–748, 2006, doi: 10.1109/TPEL.2006.872375.
- [94] J. Liu and H. Peng, "Control optimization for a power-split hybrid vehicle," in *2006 American Control Conference: Minneapolis, MN, 14 - 16 June 2006*, 2006, 6 pp, doi: 10.1109/ACC.2006.1655400.
- [95] H. A. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, and I. V. Kolmanovsky, "Predictive energy management of a power-split hybrid electric vehicle," *2009 American Control Conference, St. Louis, MO, USA*, pp. 3970–3976, 2009, doi: 10.1109/ACC.2009.5160451.

- [96] 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 Trans. Contr. Syst. Technol.*, vol. 20, no. 3, pp. 593–603, 2012, doi: 10.1109/TCST.2011.2134852.
- [97] C.-C. Lin, H. Peng, and J. W. Grizzle, "A stochastic control strategy for hybrid electric vehicles," in *Proceedings of the 2004 American Control Conference, ACC: June 30 - July 2, 2004, Boston Sheraton Hotel, Boston, Massachusetts, 2004*, 4710-4715 vol.5, doi: 10.23919/ACC.2004.1384056.
- [98] Y. Zou, Z. Kong, T. Liu, and D. Liu, "A Real-Time Markov Chain Driver Model for Tracked Vehicles and its Validation: Its Adaptability via Stochastic Dynamic Programming," *IEEE Trans. Veh. Technol.*, p. 1, 2016, doi: 10.1109/TVT.2016.2605449.
- [99] G. Ripaccioli, D. Bernardini, S. Di Cairano, A. Bemporad, and I. V. Kolmanovsky, "A stochastic model predictive control approach for series hybrid electric vehicle power management," in *Proceedings of the 2010 American Control Conference: (ACC 2010) ; Baltimore, Maryland, USA, 30 June - 2 July 2010*, 2010, pp. 5844–5849, doi: 10.1109/ACC.2010.5530504.
- [100] I. Kolmanovsky, I. Siverguina, and B. Lygoe, "Optimization of powertrain operating policy for feasibility assessment and calibration: stochastic dynamic programming approach," in *Proceedings of the 2002 American Control Conference, ACC: May 8 - 10, 2002, Hilton Anchorage and Egan Convention Center, Anchorage, Alaska, USA, 2002*, 1425-1430 vol.2, doi: 10.1109/ACC.2002.1023221.
- [101] X. Zeng and J. Wang, "Stochastic optimal control for hybrid electric vehicles running on fixed routes," in *2015 American Control Conference (ACC 2015): Chicago, Illinois, USA, 1 - 3 July 2015*, 2015, pp. 3273–3278, doi: 10.1109/ACC.2015.7171837.
- [102] M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming* (Wiley-Interscience paperback series). Hoboken, NJ: Wiley-Interscience, 2005.
- [103] M. Dyer and L. Stougie, "Computational complexity of stochastic programming problems," *Math. Program.*, vol. 106, no. 3, pp. 423–432, 2006, doi: 10.1007/s10107-005-0597-0.
- [104] H. Banvait, J. Hu, and Y. Chen, "Design of energy management system of Plug-in Hybrid Electric Vehicle using hybrid systems," in *American Control Conference (ACC), 2015: 1 - 3 July 2015, Chicago, IL, USA, 2015*, pp. 1339–1344, doi: 10.1109/ACC.2015.7170919.
- [105] E. D. Tate and S. P. Boyd, "Finding Ultimate Limits of Performance for Hybrid Electric Vehicles," *SAE Technical Paper 2000-01-3099*, vol. 2000, pp. 1–14, 2000, doi: 10.4271/2000-01-3099.

- [106] N. Murgovski, L. Johannesson, J. Sjöberg, and B. Egardt, "Component sizing of a plug-in hybrid electric powertrain via convex optimization," *Mechatronics*, vol. 22, no. 1, pp. 106–120, 2012, doi: 10.1016/j.mechatronics.2011.12.001.
- [107] N. Robuschi, M. Salazar, P. Duhr, F. Braghin, and C. H. Onder, "Minimum-fuel Engine On/Off Control for the Energy Management of a Hybrid Electric Vehicle via Iterative Linear Programming," *IFAC-PapersOnLine*, vol. 52, no. 5, pp. 134–140, 2019, doi: 10.1016/j.ifacol.2019.09.022.
- [108] S. Sager, "Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control," *Journal of Process Control*, vol. 19, no. 8, pp. 1238–1247, 2009, doi: 10.1016/j.jprocont.2009.03.008.
- [109] D.J.N. Limebeer, G. Perantoni, and A. V. Rao, "Optimal control of Formula One car energy recovery systems," *International Journal of Control*, pp. 1–16, 2014, doi: 10.1080/00207179.2014.900705.
- [110] S. Wei, Y. Zou, F. Sun, and O. Christopher, "A pseudospectral method for solving optimal control problem of a hybrid tracked vehicle," *Applied Energy*, vol. 194, pp. 588–595, 2017, doi: 10.1016/j.apenergy.2016.07.020.
- [111] J. B. Rawlings and D. Q. Mayne, *Model predictive control: Theory and design*, 1st ed. Madison, Wis.: Nob Hill Publ, 2009.
- [112] F. Yan, J. Wang, and K. Huang, "Hybrid Electric Vehicle Model Predictive Control Torque-Split Strategy Incorporating Engine Transient Characteristics," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2458–2467, 2012, doi: 10.1109/TVT.2012.2197767.
- [113] C. Sun, X. Hu, S. J. Moura, and F. Sun, "Velocity Predictors for Predictive Energy Management in Hybrid Electric Vehicles," *IEEE Trans. Contr. Syst. Technol.*, vol. 23, no. 3, pp. 1197–1204, 2015, doi: 10.1109/TCST.2014.2359176.
- [114] N. Kim, S. Cha, and H. Peng, "Optimal Control of Hybrid Electric Vehicles Based on Pontryagin's Minimum Principle," *IEEE Trans. Contr. Syst. Technol.*, vol. 19, no. 5, pp. 1279–1287, 2011, doi: 10.1109/TCST.2010.2061232.
- [115] C. Musardo, G. Rizzoni, and B. Staccia, "A-ECMS: An Adaptive Algorithm for Hybrid Electric Vehicle Energy Management," in *2005 44th IEEE Conference on Decision and Control & European Control Conference: Seville, Spain, 12 - 15 December 2005*, 2005, pp. 1816–1823, doi: 10.1109/CDC.2005.1582424.
- [116] B. Gu and G. Rizzoni, "An Adaptive Algorithm for Hybrid Electric Vehicle Energy Management Based on Driving Pattern Recognition," in *Proceedings of 2006 ASME International*

Mechanical Engineering Congress & Exposition: IMECE 2006 ; November 5 - 10, 2006, Chicago, Illinois, USA, 2006, pp. 249–258, doi: 10.1115/IMECE2006-13951.

- [117] A. Chasse, A. Sciarretta, and J. Chauvin, "Online optimal control of a parallel hybrid with costate adaptation rule," *IFAC Proceedings Volumes*, vol. 43, no. 7, pp. 99–104, 2010, doi: 10.3182/20100712-3-DE-2013.00134.
- [118] J. Hu, C. Sun, J. Xiao, and J. Li, "A torque compensation strategy in two-speed automated mechanical transmission shift process for pure electric vehicles," *Advances in Mechanical Engineering*, vol. 7, no. 11, 168781401561691, 2015, doi: 10.1177/1687814015616915.
- [119] B. Zhu *et al.*, "Gear shift schedule design for multi-speed pure electric vehicles," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 229, no. 1, pp. 70–82, 2015, doi: 10.1177/0954407014521395.
- [120] L. Guo, G. Li, B. Gao, and H. Chen, "Shift schedule optimization of 2-speed electric vehicle using model predictive control," in *Control Conference (CCC), 2014 33rd Chinese*, 2014, pp. 156–161, doi: 10.1109/ChiCC.2014.6896614.
- [121] C. T. Nguyen, P. D. Walker, S. Zhou, and N. Zhang, "Optimal sizing and energy management of an electric vehicle powertrain equipped with two motors and multi-gear ratios," *Mechanism and Machine Theory*, vol. 167, p. 104513, 2022, doi: 10.1016/j.mechmachtheory.2021.104513.
- [122] M. S. R. Mousavi, A. Pakniyat, and B. Boulet, "Dynamic modeling and controller design for a seamless two-speed transmission for electric vehicles," in *2014 IEEE Conference on Control Applications (CCA 2014): Antibes/Nice, France, 8 - 10 October 2014 ; [part of 2014 IEEE Multi-Conference on Systems and Control]*, 2014, pp. 635–640, doi: 10.1109/CCA.2014.6981411.
- [123] M. S. Rahimi Mousavi, A. Pakniyat, T. Wang, and B. Boulet, "Seamless dual brake transmission for electric vehicles: Design, control and experiment," *Mechanism and Machine Theory*, vol. 94, pp. 96–118, 2015, doi: 10.1016/j.mechmachtheory.2015.08.003.
- [124] T. Nüesch *et al.*, "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, doi: 10.1016/j.conengprac.2014.01.020.
- [125] F. Millo, L. Rolando, F. Mallamo, and R. Fuso, "Development of an optimal strategy for the energy management of a range-extended electric vehicle with additional noise, vibration and harshness constraints," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 227, no. 1, pp. 4–16, 2013, doi: 10.1177/0954407012457488.

- [126] M. Aliramezani, M. Khademnahvi, and M. Delkosh, "Optimal energy management strategy of a hybrid electric vehicle considering engine noise," *Journal of Vibration and Control*, vol. 24, no. 23, pp. 5546–5555, 2018, doi: 10.1177/1077546318758118.
- [127] J. Li, T. Huber, and C. Beidl, "Predictive Multi-Objective Operation Strategy Considering Battery Cycle Aging for Hybrid Electric Vehicles," *SAE Int. J. Alt. Power.*, vol. 7, no. 3, pp. 217–232, 2018, doi: 10.4271/2018-01-1011.
- [128] S. Ebbesen, P. Elbert, and L. Guzzella, "Battery State-of-Health Perceptive Energy Management for Hybrid Electric Vehicles," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 2893–2900, 2012, doi: 10.1109/TVT.2012.2203836.
- [129] A. Viehmann, *Erweiterte Betriebsstrategie für dedizierte parallel-serielle Hybridantriebe zur Berücksichtigung des akustischen Komforts* (Dissertation) (Forschungsberichte Mechatronische Systeme im Maschinenbau), 2020.
- [130] R. Wang and S. M. Lukic, "Review of driving conditions prediction and driving style recognition based control algorithms for hybrid electric vehicles," in *2011 IEEE Vehicle Power and Propulsion Conference (VPPC 2011): Chicago, Illinois, USA, 6 - 9 September 2011*, 2011, pp. 1–7, doi: 10.1109/VPPC.2011.6043061.
- [131] C. M. Bishop, *Pattern recognition and machine learning* (Computer science). New York, NY: Springer, 2006.
- [132] W. A. Woodward, B. P. Sadler, and S. D. Robertson, *Time series for data science: Analysis and forecasting* (Texts in statistical science series). Boca Raton, London, New York: CRC Press, 2022.
- [133] N. K. Ahmed, A. F. Atiya, N. E. Gayar, and H. El-Shishiny, "An Empirical Comparison of Machine Learning Models for Time Series Forecasting," *Econometric Reviews*, vol. 29, 5-6, pp. 594–621, 2010, doi: 10.1080/07474938.2010.481556.
- [134] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PloS one*, early access. doi: 10.1371/journal.pone.0194889.
- [135] Y. Borges Lima Ramos *et al.*, "A Comparative Study of Methods for Driver Acceleration Intention Prediction," Advanced Research Project, Institut for Mechatronic Systems, Technischen Universität Darmstadt, Darmstadt, 2020.
- [136] A. Vaswani *et al.* "Attention Is All You Need." [Online]. Available: <https://doi.org/10.48550/arXiv.1706.03762>
- [137] A. Ng. "Nuts and Bolts of Applying Deep Learning." [Online]. Available: <https://>

www.youtube.com/watch?v=F1ka6a13S9I&list=PLzIS8claVueXV_zQpsIJ06Iy1y9UwYCX2

- [138] S. Idler, P. Gwinner, K. Stahl, R. König, and S. Rinderknecht, "Innovative Super-High Multiple Speed Concept for the Electrified Automotive Powertrain," in *5th Conference on Future Automotive Technology*, Fürstenfeld, Markus Lienkamp, Ed., 2016, pp. 1–10.
- [139] M. Mileti, P. Strobl, H. Pflaum, and K. Stahl, "Design of a Hyper-High-Speed Powertrain for EV to Achieve Maximum Ranges," in *Cti Symposium 2018: 17th International Congress and Expo 3 - 6 December 2018, Berlin, Germany*, EUROFORUM Deutschland GmbH, Ed., 2020, pp. 265–273, doi: 10.1007/978-3-662-58866-6_21.
- [140] D. Schöneberger, M. Mileti, K. Stahl, and S. Rinderknecht, "Development of an Innovative Shift Actuator for Electrified Multispeed Transmissions," *2019 International Conference on Advanced Vehicle Powertrains*, 2019.
- [141] D. Schöneberger, S. Rinderknecht, D. Reitmeier, and A. Mertens, "Development of a Transmission Control for an Innovative High-speed Powertrain using Motor-related Control: Simulative approach to develop an angular position controlled engagement of dog clutches in a two drive transmission," *Drivetrain for Vehicles 2020*, I-195-I-212, 2020, doi: 10.51202/9783181023730-I-195.
- [142] D. Schöneberger, *Entwicklung eines dedizierten Schaltungssystems für elektrische Mehrgang-Fahrzeugantriebe* (Dissertation), 1st ed. (Forschungsberichte Mechatronische Systeme im Maschinenbau). Düren: Shaker Verlag, 2022.
- [143] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Math. Comp.*, vol. 19, no. 92, pp. 577–593, 1965, doi: 10.1090/S0025-5718-1965-0198670-6.
- [144] T. N. E. Greville, "Note on the Generalized Inverse of a Matrix Product," *SIAM Rev.*, vol. 8, no. 4, pp. 518–521, 1966, doi: 10.1137/1008107.
- [145] C. Y. n. Kaya, S. K. Lucas, and S. T. Simakov, "Computations for bang-bang constrained optimal control using a mathematical programming formulation," *Optim. Control Appl. Meth.*, vol. 25, no. 6, pp. 295–308, 2004, doi: 10.1002/oca.749.
- [146] D. P. Bertsekas, *Convex optimization algorithms* (Optimization and computation series 4). Belmont, Massachusetts: Athena Scientific, 2015.
- [147] C. Cartis, N. I.M. Gould, and P. L. Toint, "On the complexity of finding first-order critical points in constrained nonlinear optimization," *Math. Program.*, vol. 144, 1-2, pp. 93–106, 2014, doi: 10.1007/s10107-012-0617-9.
- [148] F. E. Curtis, M. J. O'Neill, and D. P. Robinson, "Worst-case complexity of an SQP method for nonlinear equality constrained stochastic optimization," *Math. Program.*, 2023, doi:

10.1007/s10107-023-01981-1.

- [149] F. A. Potra and S. J. Wright, "Interior-point methods," *Journal of Computational and Applied Mathematics*, vol. 124, 1-2, pp. 281–302, 2000, doi: 10.1016/S0377-0427(00)00433-7.
- [150] The MathWorks Inc. "MATLAB version: 9.13.0 (R2022b)." [Online]. Available: <https://de.mathworks.com/help/optim/ug/tolerances-and-stopping-criteria.html>
- [151] M. J. D. Powell, "A fast algorithm for nonlinearly constrained optimization calculations," in *Numerical analysis: Proceedings of the biennial conference, held at Dundee, June 28-July 1, 1977*, G. A. Watson, Ed., vol. 630, 1978, pp. 144–157, doi: 10.1007/BFb0067703.
- [152] M.J.D. Powell, "THE CONVERGENCE OF VARIABLE METRIC METHODS FOR NONLINEARLY CONSTRAINED OPTIMIZATION CALCULATIONS," in *Nonlinear programming 3: Proceedings of the Special Interest Group on Mathematical Programming symposium*, O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, Eds., 1978, pp. 27–63, doi: 10.1016/B978-0-12-468660-1.50007-4.
- [153] B. Bertsche, *Reliability in automotive and mechanical engineering: Determination of component and system reliability*. Berlin, Heidelberg: Springer, 2008.
- [154] L. U. Gokdere, S. L. Chiu, K. J. Keller, and J. Vian, "Lifetime control of electromechanical actuators," *IEEE Aerospace Conference Proceedings*, pp. 3523–3531, 2005, doi: 10.1109/AERO.2005.1559655.
- [155] M. F. Cabanas *et al.*, Eds., *Analysis of the fatigue causes on the rotor bars of squirrel cage asynchronous motors: experimental analysis and modelling of medium voltage motors*. IEEE, 2003, doi: 10.1109/DEMPED.2003.1234581.
- [156] M. A. Laughton, *Electrical Engineer's Reference Book*, 16th ed. (Electrical Engineer's Reference Book). Kidlington: Elsevier Science & Technology, 2003.
- [157] E. L. Brancato, "Estimation of lifetime expectancies of motors," *IEEE Electr. Insul. Mag.*, vol. 8, no. 3, pp. 5–13, 1992, doi: 10.1109/57.139066.
- [158] V. M. Montsinger, "Loading Transformers By Temperature," *Trans. Am. Inst. Electr. Eng.*, vol. 49, no. 2, pp. 776–790, 1930, doi: 10.1109/T-AIEE.1930.5055572.
- [159] T. W. Dakin, "Electrical Insulation Deterioration Treated as a Chemical Rate Phenomenon," *Trans. Am. Inst. Electr. Eng.*, vol. 67, no. 1, pp. 113–122, 1948, doi: 10.1109/T-AIEE.1948.5059649.
- [160] Y. Tao, J. Li, G. Gao, Z. Liu, and S. Rinderknecht, "Goal-Oriented Data-Driven Control for a

Holistic Thermal Management System of an Electric Vehicle," *IEEE Trans. Automat. Sci. Eng.*, pp. 1–12, 2023, doi: 10.1109/TASE.2023.3304521.

- [161] I. Boiadjiev, J. Witzig, T. Tobie, and K. Stahl, "Tooth flank fracture – basic principles and calculation model for a sub surface initiated fatigue failure mode of case hardened gears," in *International Gear Conference 2014: 26th-28th August 2014, Lyon Villeurbanne, France*, 2014, pp. 670–680, doi: 10.1533/9781782421955.670.
- [162] G. Jacobs, M. Plogmann, and T. Mang, Eds. *Rolling Bearing Damages // Encyclopedia of lubricants and lubrication*. Berlin: Springer, 2014.
- [163] G. Doll, *Rolling Bearing Tribology: Tribology and Failure Modes of Rolling Element Bearings*. San Diego: Elsevier, 2023.
- [164] A. Grabulov, *Fundamentals of rolling contact fatigue* (Zugl.: Delft University of Technology, Diss., 2010). [S.l.]: [s.n.], 2010.
- [165] M. Wacker, *Einfluss von Drehungleichförmigkeiten auf die Zahnradlebensdauer in Fahrzeuggetrieben* (Zugl.: Stuttgart, Univ., Diss., 2013 (Nicht für den Austausch)) (Berichte aus dem Institut für Maschinenelemente, Antriebstechnik, CAD, Dichtungen, Zuverlässigkeit 144). Stuttgart: IMA, 2013.
- [166] S. Foulard, *Online and real-time load monitoring for remaining service life prediction of automotive transmissions: Damage level estimation of transmission components based on a torque acquisition* (Zugl.: Darmstadt, Techn. Univ., Diss., 2015 / Zugl.: Lyon, Ecole Centrale, Diss., 2015) (Forschungsberichte Mechatronische Systeme im Maschinenbau). Aachen: Shaker, 2015.
- [167] G. Lundberg and A. Palmgren, "Dynamic Capacity of Rolling Bearings," *Journal of Applied Mechanics*, vol. 16, no. 2, pp. 165–172, 1949, doi: 10.1115/1.4009930.
- [168] E. Haibach, *Betriebsfestigkeit: Verfahren und Daten zur Bauteilberechnung*, 3rd ed. (VDI-Buch). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2006.
- [169] J.-I. Hwang and G. Poll, "A new approach for the prediction of fatigue life in rolling bearings based on damage accumulation theory considering residual stresses," *Front. Manuf. Technol.*, vol. 2, 2022, Art. no. 1010759, doi: 10.3389/fmtec.2022.1010759.
- [170] M. A. Miner, "Cumulative Damage in Fatigue," *Journal of Applied Mechanics*, vol. 12, no. 3, A159-A164, 1945, doi: 10.1115/1.4009458.
- [171] M. Mistler, *Lebensdauerprognose für dynamisch beanspruchte Elastomerbauteile auf Basis der thermo-mechanischen Materialbeanspruchung*. Duisburg: DuEPublico, 2018. [Online]. Available: <https://nbn-resolving.org/urn:nbn:de:hbz:464-20180629-151910-4>

- [172] A. G. Palmgren, "Die Lebensdauer von Kugellagern," *Zeitschrift des Vereines Deutscher Ingenieure (ZVDI)*, vol. 1924, no. 68, 1924, Art. no. 14.
- [173] Tauscher H. and Buchholz H., "Dauerschwingverhalten der Stähle unter extrem langer Beanspruchungsdauer," *Maschinenbautechnik*, vol. 1974, no. 12, pp. 565–567, 1974.
- [174] Bathias, "There is no infinite fatigue life in metallic materials," *Fatigue & Fracture of Engineering Materials & Structures*, vol. 22, no. 7, pp. 559–565, 1999, doi: 10.1046/j.1460-2695.1999.00183.x.
- [175] E. Haibach, "Modifizierte lineare Schadensakkumulationshypothese zur Berücksichtigung des Dauerfestigkeitsabfalls mit fortschreitender Schädigung," *Laboratorium für Betriebsfestigkeit*, Darmstadt, Rep. 50/70, 1970.
- [176] J. Liu and H. Zenner, "Berechnung von Bauteilwöhlerlinien unter Berücksichtigung der statistischen und spannungsmechanischen Stützziffer," *Mat.-wiss. u. Werkstofftech.*, vol. 26, no. 1, pp. 14–21, 1995, doi: 10.1002/mawe.19950260106.
- [177] Sculley, D., Snoek, J., Wiltschko, A.B., & Rahimi, A., "Winner's Curse? On Pace, Progress, and Empirical Rigor," in *International Conference on Learning Representations: 30th April - 3rd May, 2018, Vancouver, BC, Canada*, 2018, pp. 1–4.

Standards

- [1] *IEEE 754: "IEEE Standard for Floating-Point Arithmetic,"* IEEE Std 754-2008, vol., no., pp.1-70, 29 Aug. 2008, doi: 10.1109/IEEESTD.2008.4610935.
- [2] *DIN 3990: Calculation of load capacity of cylindrical gears: Calculation of service strength.* 1994. DIN Deutsches Institut für Normung e.V., Berlin.
- [3] *DIN ISO 281: Rolling Bearings – Dynamic Load Ratings and Rating Life.* 2003. DIN Deutsches Institut für Normung e.V., Berlin.
- [4] *DIN 50100: Load controlled fatigue testing - Execution and evaluation of cyclic tests at constant load amplitudes on metallic specimens and components.* 2016. DIN Deutsches Institut für Normung e.V., Berlin.