



# **Optimising Power Management Strategies for Railway Traction Systems**

by

**SHAOFENG LU**

A thesis submitted to

The University of Birmingham

for the degree of

**DOCTOR OF PHILOSOPHY**

School of Electronic, Electrical and Computer Engineering

College of Engineering and Physical Sciences

The University of Birmingham, UK

October 2011

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

# Abstract

Railway transportation is facing increasing pressure to reduce the energy demand of its vehicles due to increasing concern for environmental issues. This thesis presents studies based on improved power management strategies for railway traction systems and demonstrates that there is potential for improvements in the total system energy efficiency if optimised high-level supervisory power management strategies are applied. Optimised power management strategies utilise existing power systems in a more cooperative and energy-efficient manner in order to reduce the total energy demand. In this thesis, three case studies in different research scenarios have been conducted.

Under certain operational, geographic and physical constraints, the energy consumed by the train can be significantly reduced if improved control strategies are implemented. This thesis proposes a distance based model for train speed trajectory optimisation. Three optimisation algorithms, Ant Colony Optimisation (ACO), Genetic Algorithm (GA) and Dynamic Programming (DP), are applied to search for the optimal train speed trajectory, given a journey time constraint. The speed at each preset position along the journey is determined and optimised using these search algorithms.

In a DC railway network, power peaks in a substation are not desirable as they could present safety risks and are not energy efficient. A power peak can be avoided if the control of multiple trains is coordinated. The allocation of inter-station journey time intrinsically affects both service quality and energy efficiency. By identifying an optimal journey time allocation, a multi-objective function targeting both

---

energy efficiency and service quality can be used. In this study, a DC railway is modelled with two parallel railway tracks, five station stops and three DC electric substations. Regenerative braking is studied in this DC electric network using Nodal Analysis (NA) and the Load Flow (LF) method. This study demonstrates that within the neighbourhood of an electric railway network, the journey time allocation for inter-station journeys will affect the total service quality and the energy loss. A GA is applied to find the best inter-station journey time allocation.

Finally, this thesis explores the potential of applying advanced power management strategies to a Diesel Multiple Unit (DMU) train. DMU trains have multiple diesel engines which are commonly operated in a homogenous manner. The work presented in this thesis analyses the potential energy savings that may be obtained through the independent operation of the engines. Two widely investigated power management strategies which have been applied to the control of Hybrid Electric Vehicles are studied for a typical DMU railway vehicle. DP is applied to identify the optimal instant power distribution between engines. Based on the optimised results from DP, an adaptive rule-based online strategy is proposed using a non-linear programming optimisation algorithm.

# Acknowledgements

I am heartily thankful to Dr. Stuart Hillmansen. Without his encouragement, inspiration, support and supervision, this thesis would not be possible. I would also like to extend my deep gratitude to Prof. Clive Roberts for his invaluable guidance and support which has enable me to benefit most out of the research.

I would like to sincerely thank my wife Ms. Hui Zhou who has been there to love, support and encourage me with great patience and understanding during my PhD. I am grateful to my family who have always loved and tolerated me. I would like to thank Mr. Yudong Wu, Dr. Paul Western, Ms Sharon Berry and Ms. Katherine Slate for their help.

I am grateful to all the people who I am not able to list but have helped and supported me over the past four years.

# Table of Contents

|   |             |
|---|-------------|
| <b>Table of Contents</b>                      | <b>iv</b>   |
| <b>List of Figures</b>                        | <b>ix</b>   |
| <b>List of Tables</b>                         | <b>xiii</b> |
| <b>List of Acronyms</b>                       | <b>xiv</b>  |
| <b>1 Introduction</b>                         | <b>1</b>    |
| 1.1 Background . . . . .                      | 1           |
| 1.2 Case studies overview . . . . .           | 2           |
| 1.3 Objectives . . . . .                      | 4           |
| 1.4 Thesis structure . . . . .                | 5           |
| <b>2 A review of railway traction systems</b> | <b>8</b>    |
| 2.1 Introduction . . . . .                    | 8           |
| 2.2 Railway traction systems . . . . .        | 9           |
| 2.2.1 Electric traction . . . . .             | 9           |
| 2.2.2 Diesel-electric traction . . . . .      | 10          |
| 2.2.3 Hybrid traction . . . . .               | 11          |
| 2.3 DC motor drive . . . . .                  | 12          |
| 2.4 AC motor drive . . . . .                  | 14          |
| 2.5 Summary . . . . .                         | 19          |

---

## TABLE OF CONTENTS

---

|          |  |           |
|----------|--|-----------|
| <b>3</b> | <b>Review of optimisation techniques</b>               | <b>21</b> |
| 3.1      | Introduction . . . . .                                 | 21        |
| 3.2      | Numerical Optimisation . . . . .                       | 22        |
| 3.3      | Dynamic programming . . . . .                          | 26        |
| 3.3.1    | Introduction . . . . .                                 | 26        |
| 3.3.2    | Mathematical presentation . . . . .                    | 27        |
| 3.3.3    | Elements of dynamic programming . . . . .              | 30        |
| 3.3.4    | Dynamic programming and the greedy algorithm . . . . . | 32        |
| 3.3.5    | Summary . . . . .                                      | 32        |
| 3.4      | Metaheuristics . . . . .                               | 33        |
| 3.4.1    | Introduction . . . . .                                 | 33        |
| 3.4.2    | Genetic algorithm . . . . .                            | 35        |
| 3.4.3    | Ant Colony Optimisation (ACO) . . . . .                | 36        |
| 3.4.4    | Summary . . . . .                                      | 39        |
| 3.5      | Summary . . . . .                                      | 40        |
| <b>4</b> | <b>Modelling train motion and traction power</b>       | <b>41</b> |
| 4.1      | Introduction . . . . .                                 | 41        |
| 4.2      | Physics of vehicle motion . . . . .                    | 42        |
| 4.2.1    | General introduction . . . . .                         | 42        |
| 4.2.2    | Adhesion . . . . .                                     | 43        |
| 4.2.3    | Resistance . . . . .                                   | 45        |
| 4.2.4    | Effective Mass . . . . .                               | 46        |
| 4.2.5    | General vehicle motion equation . . . . .              | 47        |
| 4.3      | Modelling and simulation . . . . .                     | 47        |
| 4.3.1    | Vehicle state switch . . . . .                         | 48        |
| 4.3.2    | Operational control input . . . . .                    | 50        |
| 4.3.3    | Energy consumption modelling . . . . .                 | 56        |
| 4.3.4    | Single train motion simulator . . . . .                | 57        |

---

---

## TABLE OF CONTENTS

---

|          |  |            |
|----------|--|------------|
| 4.4      | Summary . . . . .  | 59         |
| <b>5</b> | <b>Single train trajectory optimisation</b>  | <b>60</b>  |
| 5.1      | Introduction . . . . .   | 60         |
| 5.2      | An optimal control view-point . . . . .  | 62         |
| 5.2.1    | System plant . . . . .   | 62         |
| 5.2.2    | The pontryagin minimum principle . . . . .   | 63         |
| 5.2.3    | Singular control solution . . . . .  | 65         |
| 5.3      | Modelling context . . . . .  | 69         |
| 5.3.1    | Introduction . . . . .   | 69         |
| 5.3.2    | Speed limitation . . . . .   | 71         |
| 5.3.3    | Minor speed switch . . . . .   | 73         |
| 5.3.4    | Sparse data storage . . . . .  | 75         |
| 5.4      | ACO application . . . . .  | 75         |
| 5.4.1    | Introduction . . . . .   | 75         |
| 5.4.2    | Construction graph . . . . .   | 76         |
| 5.4.3    | Solution construction . . . . .  | 78         |
| 5.4.4    | Pheromone update and termination condition . . . . .                               | 81         |
| 5.5      | Genetic algorithm . . . . .  | 83         |
| 5.5.1    | Introduction . . . . .   | 83         |
| 5.5.2    | Genotype generation . . . . .  | 84         |
| 5.6      | Dynamic programming . . . . .  | 86         |
| 5.6.1    | Introduction . . . . .   | 86         |
| 5.6.2    | Optimisation process . . . . .   | 87         |
| 5.6.3    | Summary . . . . .  | 95         |
| 5.7      | Results and discussion . . . . .   | 96         |
| 5.8      | Summary . . . . .  | 102        |
| <b>6</b> | <b>Optimising the coordinated train operation in a DC railway electric network</b> | <b>103</b> |

---



---

## TABLE OF CONTENTS

---

|          |   |            |
|----------|---|------------|
| 6.1      | Introduction . . . . .  | 103        |
| 6.2      | DC railway network modelling . . . . .                              | 105        |
| 6.2.1    | DC electrical railway network . . . . .                             | 105        |
| 6.2.2    | DC electrical analysis . . . . .                                    | 107        |
| 6.2.3    | Single train simulation using coasting control . . . . .            | 108        |
| 6.2.4    | Iterative power flow calculation . . . . .                          | 110        |
| 6.3      | Methodology . . . . .   | 114        |
| 6.3.1    | Objective function . . . . .  | 114        |
| 6.3.2    | Journey time allocation . . . . .                                   | 116        |
| 6.3.3    | The genetic algorithm . . . . .                                     | 117        |
| 6.4      | Results and discussion . . . . .                                    | 119        |
| 6.5      | Summary . . . . .   | 123        |
| <b>7</b> | <b>A power management strategy for a Diesel Multiple Unit train</b> | <b>125</b> |
| 7.1      | Review of power management strategies . . . . .                     | 126        |
| 7.2      | Typical DMU train . . . . .   | 128        |
| 7.3      | Engine description . . . . .  | 129        |
| 7.3.1    | Engine efficiency map . . . . .                                     | 129        |
| 7.3.2    | Diesel energy consumed calculation . . . . .                        | 130        |
| 7.4      | Problem definition . . . . .  | 131        |
| 7.4.1    | Objective and constraints . . . . .                                 | 131        |
| 7.4.2    | Total power state vector and diesel fuel cost . . . . .             | 133        |
| 7.5      | Solutions and Results . . . . .                                     | 136        |
| 7.5.1    | Dynamic Programming . . . . .                                       | 137        |
| 7.5.2    | Adaptive online strategies . . . . .                                | 141        |
| 7.5.3    | Results . . . . .   | 150        |
| 7.6      | Summary . . . . .   | 152        |
| <b>8</b> | <b>Conclusions and future work</b>                                  | <b>154</b> |
| 8.1      | General summary of contents . . . . .                               | 154        |

---

---

## TABLE OF CONTENTS

---

|          |   |            |
|----------|---|------------|
| 8.2      | Conclusions . . . . .                                       | 155        |
| 8.2.1    | Single train trajectory optimisation study . . . . .        | 155        |
| 8.2.2    | Journey time allocation study . . . . .                     | 157        |
| 8.2.3    | A power management strategy for a multiple unit train . .   | 158        |
| 8.3      | Future work . . . . .                                       | 160        |
| 8.3.1    | Validation and verification . . . . .                       | 160        |
| 8.3.2    | Single train trajectory optimisation study . . . . .        | 160        |
| 8.3.3    | Journey time allocation study . . . . .                     | 160        |
| 8.3.4    | Power management strategies study for a multiple unit train | 161        |
| <b>A</b> | <b>Electric drive for railway traction systems</b>          | <b>162</b> |
| A.1      | DC motor drive . . . . .                                    | 162        |
| A.1.1    | DC-DC chopper converter traction drive . . . . .            | 162        |
| A.1.2    | Phase-controlled rectifier traction drive . . . . .         | 165        |
| A.2      | AC motor drive . . . . .                                    | 167        |
| A.2.1    | DC-fed Current Source Inverter drive . . . . .              | 167        |
| A.2.2    | DC-fed Voltage Source Inverter drive . . . . .              | 168        |
| A.2.3    | AC-fed Voltage Source Inverter drive . . . . .              | 170        |
| <b>B</b> | <b>Optimisation Techniques</b>                              | <b>173</b> |
| B.1      | Categories of Numerical Optimisation . . . . .              | 173        |
| B.2      | Some other metaheuristics . . . . .                         | 175        |
| B.2.1    | Simulated annealing . . . . .                               | 175        |
| B.2.2    | Tabu Search (TS) . . . . .                                  | 178        |
| <b>C</b> | <b>Publications</b>   | <b>181</b> |
|          | <b>References</b>   | <b>182</b> |

# List of Figures

|     |   |    |
|-----|---|----|
| 1.1 | Categories of energy efficiency improvement technologies . . . .                        | 3  |
| 2.1 | An architecture diagram for a diesel-electric propulsion system . .                     | 11 |
| 2.2 | An architecture diagram for a diesel hybrid propulsion system. . .                      | 13 |
| 2.3 | Separately excited DC motor traction drive regime of operation . .                      | 14 |
| 2.4 | Typical induction motor torque-speed characteristic curve . . . .                       | 16 |
| 2.5 | Mechanical and electrical variables over traction duty cycle . . . .                    | 18 |
| 4.1 | Typical railway vehicle tractive effort and resistance characteristic                   | 44 |
| 4.2 | Vehicle state switch between two states in 3 dimension space . . .                      | 49 |
| 4.3 | Vehicle states switch in 3 dimension space . . . . .                                    | 51 |
| 4.4 | Maximum achievable acceleration truncation using partial open loop<br>control . . . . . | 52 |
| 4.5 | Coasting mode operational control . . . . .   | 53 |
| 4.6 | Braking mode operational control due to reduced speed limit . . .                       | 54 |
| 4.7 | Braking mode operational control due to station stop . . . . .                          | 55 |
| 4.8 | Braking mode operational control at distance based modelling . .                        | 56 |
| 4.9 | System diagram of single train simulator . . . . .                                      | 58 |
| 5.1 | Speed selection procedure in the distance based trajectory searching                    | 70 |
| 5.2 | Various typical trajectories between two pre-determined positions                       | 74 |
| 5.3 | Demonstration of the speed link matrix . . . . .  | 76 |
| 5.4 | Various speed change signal for one pre-determined position interval                    | 85 |

---

## LIST OF FIGURES

---

|      |   |     |
|------|---|-----|
| 5.5  | Schematic Genetic Algorithm optimisation of train speed trajectory.   | 87  |
| 5.6  | Vehicle states generation procedure in DP. . . . .  | 89  |
| 5.7  | Admissible area of train states in DP search procedure . . . . .  | 95  |
| 5.8  | Maximum tractive effort, resistance and acceleration curve of Voyager type vehicle . . . . .  | 96  |
| 5.9  | Optimised journey trajectories using ACO under different journey time conditions . . . . .  | 97  |
| 5.10 | Optimised journey trajectories using GA under different journey time conditions . . . . .   | 98  |
| 5.11 | Optimised journey trajectories using DP under different journey time conditions . . . . .   | 99  |
| 5.12 | The best and mean objective function output for a journey time of 2800 s at each generation for ACO . . . . .                                       | 99  |
| 5.13 | The best and mean objective function output for a journey time of 2800 s at each generation for GA . . . . .  | 100 |
| 5.14 | Journey energy cost vs. time cost curves using different algorithms   | 100 |
| 6.1  | A simplified DC railway electrical network diagram . . . . .  | 106 |
| 6.2  | Physical location arrangement of stations and substations in a DC railway network . . . . .   | 106 |
| 6.3  | Nodal analysis . . . . .  | 107 |
| 6.4  | Maximum tractive effort and resistance curve of a typical urban railway vehicle. . . . .  | 109 |
| 6.5  | Vehicle state switch using coasting control. . . . .  | 109 |
| 6.6  | Optimised train running trajectory and instant power output using coasting control strategies for journey between station 1 and station 2 . . . . . | 110 |
| 6.7  | Non linear relationship between the terminal voltage and current of a train . . . . .   | 112 |

---

---

## LIST OF FIGURES

---

|      |  |     |
|------|--|-----|
| 6.8  | Non linear relationship between the terminal voltage and injected current out of substation. . . . .                       | 113 |
| 6.9  | Genetic algorithm implementation chart flow . . . . .  | 118 |
| 6.10 | Entire optimisation procedure using Genetic Algorithm. . . . .   | 119 |
| 6.11 | Train speed vs. time . . . . .   | 120 |
| 6.12 | Train distance vs. time . . . . .  | 121 |
| 6.13 | Instant voltage output across a headway time 240s. . . . .   | 122 |
| 6.14 | Instant current output across a headway time 240s. . . . .   | 122 |
| 6.15 | Evolutionary fitness function output vs. generation in the GA . . .  | 123 |
| 6.16 | Distance vs. Time for forward journeys with different journey time allocation . . . . .                                    | 124 |
| 7.1  | A typical 3 coach DMU train . . . . .  | 128 |
| 7.2  | Controlled diesel engine power efficiency characteristic . . . . .   | 130 |
| 7.3  | Engine power states approximation . . . . .  | 133 |
| 7.4  | Total power state vector in 3-D spaces; the magnitude of the vector is 100 at any point at the triangular surface. . . . . | 135 |
| 7.5  | Relationship between engine states and fuel cost under various power demands . . . . .                                     | 136 |
| 7.6  | Dynamic programming evolution diagram . . . . .  | 139 |
| 7.7  | Preferable number of engines for various power demands . . . . .   | 143 |
| 7.8  | Illustration for minimisation of discrepancy . . . . .   | 144 |
| 7.9  | Speed profile and speed limit . . . . .  | 146 |
| 7.10 | Distance time and distance elevation graph . . . . .   | 146 |
| 7.11 | Power distribution over time using dynamic programming and on-line rule based strategy (separate version) . . . . .        | 147 |
| 7.12 | Power distribution using DP using both online and rule based strategy  | 148 |
| 7.13 | Initial stage of the power distribution for both power management strategies . . . . .                                     | 149 |

---

---

## LIST OF FIGURES

---

|  |     |
|--|-----|
| 7.14 Power distribution by dynamic programming for two further case studies . . . . .              | 150 |
| A.1 Two-quadrant traction chopper power operational circuit and the operational waveform . . . . . | 163 |
| A.2 Chopper power circuit corresponding to state 1 and 4 in Figure A.1(a)                          | 164 |
| A.3 Chopper power circuit corresponding to state 3 and 2 in Figure A.1(a)                          | 165 |
| A.4 Semi-controlled single-phase bridge operation circuit and waveforms                            | 166 |
| A.5 Traction drives with three-phase motors and DC power supply . .                                | 168 |
| A.6 Current source inverter circuit diagram . . . . .  | 169 |
| A.7 AC power supply with induction motor drive . . . . .   | 170 |
| A.8 Circuit diagram of a pulse converter with AC input and DC output.                              | 171 |
| A.9 Phase diagram of AC-DC pulse converter . . . . .   | 171 |
| B.1 The classification of the mathematical programming problem . . .                               | 175 |

# List of Tables

|     |  |     |
|-----|--|-----|
| 5.1 | Rules to decide heuristic coefficients . . . . .   | 81  |
| 5.2 | Next speed selection based on the characterised control index number   | 85  |
| 5.3 | Key parameters for Single Train Motion Simulator . . . . .   | 96  |
| 5.4 | Computational time comparison between three algorithms applied<br>on the journey with time requirement of 2800 s for 15 runs . . . . | 101 |
| 7.1 | Typical technical data for DMU type vehicles . . . . .   | 129 |
| 7.2 | Threshold power demand value . . . . .   | 142 |
| 7.3 | Preferable power distribution based on various power demand . .  | 143 |
| 7.4 | Online power distribution . . . . .  | 145 |
| 7.5 | Simulation results comparison . . . . .  | 152 |

# List of Acronyms

**AC** Alternating Current.

**ACO** Ant Colony Optimisation.

**ANN** Artificial Neural Network.

**ATO** Automatic Train Operation.

**BF** Brute Force.

**BR** British Rail.

**BSFC** Brake Specific Fuel Consumption.

**CBD** Critical Braking Distance.

**CSI** Current Source Inverter.

**DC** Direct Current.

**DMU** Diesel Multiple Unit.

**DP** Dynamic Programming.

**ECMS** Equivalent Consumption Minimisation Strategy.

**EMF** Electromotive Force.

**EMU** Electric Multiple Unit.



**ESD** Energy Storage Device.

**FOR** Flat Out Running.

**GA** Genetic Algorithm.

**GRASP** Greedy Randomised Adaptive Search Procedures.

**GTO** Gate-Turn-Off.

**HEV** Hybrid Electric Vehicle.

**ICE** Internal Combustion Engine.

**IGBT** Insulated Gate Bi-polar Transistor.

**IM** Induction Motor.

**LF** Load Flow.

**LSM** Least Square Minimisation.

**MPS** Maximum Power Switch.

**NA** Nodal Analysis.

**ODE** Ordinary Derivative Equation.

**PMP** Pontryagin Maximum Principle.

**PNE** Preferable Number of Engines.

**PWM** Pulse Width Modulation.

**SDP** Semidefinite programming.

**SOCP** Second-order Conic Programming.

## List of Acronyms

---

**STMS** Single Train Motion Simulator.

**TS** Tabu Search.

**VSI** Voltage Source Inverter.

# Chapter 1

## Introduction

### 1.1 Background

Rail travel is often regarded as the energy efficient and environmentally friendly mode of transport as demonstrated in [1]. Rail transport owes its excellent energy efficiency, in part, to the low friction between the wheel and rail. It is self-evident that, due to the very hard contact, little energy is dissipated through rolling resistance [2]. Railway vehicles also have good aerodynamic performance due to the benefits of the convoy formation [3].

The railway transportation industry is facing increasing pressure to further improve energy efficiency with the growing concern for environmental issues [4]. Energy used by train operations accounts for around 70% of all energy used by the railway system [5]. It follows that it is not only environmentally beneficial but also economically advantageous to reduce energy consumption. Railway transportation will not remain a viable solution for future mobility if energy efficiency cannot be improved. Two methods are proposed to improve efficiency: one is to improve technology; the other is to improve organisation and production (i.e. operations) [6].

Advancements in technology have the potential to improve the energy efficiency of almost every aspect of the railway system. For example, a reduction in train mass

has the potential to reduce energy consumption significantly [7]. Advanced train surface design can improve aerodynamic performance and hence reduce the energy consumed [3]. With the advent of modern electronic and electrical technology, such as engine control technology, power electronics and electric motor control, energy efficiency can be further improved [8–11]. Regenerative braking has been used to improve energy efficiency for railway system [12–14]. For non-electrified routes, one approach is to install energy storage devices, such as batteries, super capacitors, flywheels, etc. either trackside or on-board trains [15–17]. For electrified routes, regenerative braking power can be fed back into the power system [13, 18].

New technologies are often expensive to deploy on existing railway systems, however, changes to operational procedures, such as the application of energy efficient driving can be rapidly and cost effectively implemented. Energy efficient driving can be achieved in a number of ways, ranging from improved driver training to dynamic driver advisory systems, which use real-time optimisation approaches [19]. Advanced railway control can be used to improve traffic fluidity and hence reduce energy consumption [20, 21]. The application of the two methods to the rail industry to improve energy efficiency is shown in Figure 1.1. This thesis will focus on operational procedures for improving energy efficiency in railway systems. A number of case studies are considered where it is shown that through using control strategies, energy efficiency improvements can be made.

## 1.2 Case studies overview

The three interrelated studies considered in this thesis are:

1. The optimisation of the speed trajectory for a single train.
2. Improvements in system energy efficiency for multiple trains on an electric network.
3. Optimisation of the power sources on-board a diesel multiple-unit train.

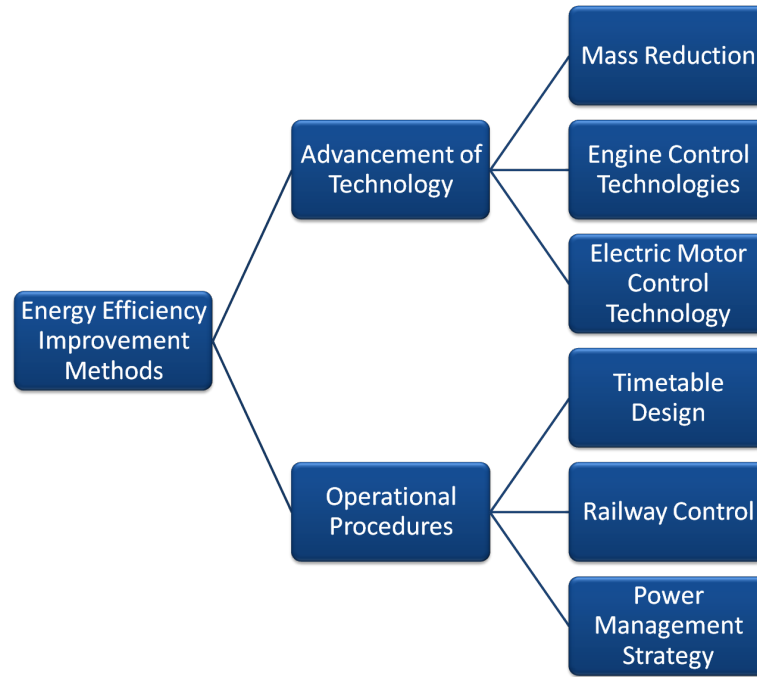


Figure 1.1: Categories of energy efficiency improvement technologies

In the first case study, a train is considered as a single system. Using various control signals, trains are able to accelerate, cruise, coast and brake to modify the train speed trajectory. Using either on-line or off-line computer optimisation algorithms, it is possible to obtain an optimal train speed trajectory. In practice, the behaviour of the driver will significantly affect energy consumption [22]. To address this issue, the optimal train speed trajectory can be provided through a driving advisory system or automatic train control.

Numerous researchers have considered how to obtain an optimal train speed trajectory. The research can be grouped into two categories: coasting control and overall control. Coasting control aims to improve the energy efficiency of a single train operation by applying coasting, i.e. the traction motors are switched off completely at appropriate points [23–27]. Non-deterministic optimisation methods, e.g. Genetic Algorithm (GA), are commonly applied in this type of study. However, since coasting is the only control signal, the generality is compromised. The second category, overall control, uses all possible control signals [19, 28–33].

Due to the generic nature of this approach, the algorithms used are generally computationally intensive. In the first case study, this thesis develops a practical overall control strategy that is able to find energy efficient single train speed trajectories.

Conventionally each train in an inter-connected Direct Current power system is considered independently. However, in practice, the trains are interrelated via the power system. System-wide power management strategies are considered in the second case study in order to reduce the average current level in the network [34–37] and improve the power system receptivity [26, 34, 38, 39]. In the second case study, using both Nodal Analysis and Load Flow method, optimal journey time distribution is obtained to reduce the power peaks in the substations. It is found that increase of part of inter-station journey time can increase total energy efficiency.

In the third case study, multiple-unit trains with independently controllable traction units are considered. If the traction units of a multiple-unit train can be controlled independently, it is considered whether energy savings can be achieved if optimised supervisory power management strategies are employed, energy saving can be achieved. Previous literatures in this area have mainly focused on power management strategies for hybrid electric vehicles [5, 40–47]. Recent research has also considered power management strategies of improving the fuel consumption for Diesel Multiple Unit (DMU) trains [48–50].

## 1.3 Objectives

The objectives of the study presented in this thesis are as follows.

- Under various constraints, the energy efficiency for the train speed trajectory can be significantly improved if proper control strategies are implemented. This thesis targets a single train as a single power source and proposes a distance-based and general control model for single speed trajectory optimisation. Three optimisation algorithms, i.e. Ant Colony Optimisation, Genetic

Algorithm and Dynamic Programming are applied to search the optimal train speed trajectory under the journey time constraint. Whilst the objective of obtaining an optimal train speed trajectory remains the main target of this case, an extensive study on the algorithms will be done comparatively to identify the characteristic of each algorithm.

- In a DC railway network, the power peaks in a substation is not desirable for the sake of both safety and energy efficiency. The power peak can be reduced if coordinative operations of trains are taken. This thesis targets on multiple trains in a DC electric network as multiple power sources and demonstrates that within a neighbourhood of an electric railway network, inter-operation between trains will affect energy efficiency. Genetic Algorithm is applied to find the optimal distribution of inter-station journey time to improve the energy efficiency of the entire electric network.
- Improved fuel consumption and lower emissions are two of the key objectives for future transportation. This thesis explores the potential of applying advanced power management strategies to a Diesel Multiple Unit train as a combination of multiple power sources. This thesis analyses the potential energy savings that may be obtained through the independent operation of the engines. Two widely investigated power management strategies are applied to a typical Diesel Multiple Unit railway vehicle. Dynamic Programming strategies will be applied to identify the optimal instant power distribution between the engines. An adaptive rule-based online strategy based on the optimisation results from the Dynamic programming solution is then proposed and realised using a non-linear programming optimisation algorithm.

## 1.4 Thesis structure

The thesis is set out as follows:

---

## 1.4. THESIS STRUCTURE

---

- Chapter 1 is a general introduction to the background and research motivations, case studies, objectives and thesis structure.
- Chapter 2 provides a review of railway traction power systems, including the power supply system, the Direct Current (DC) motor drive system and the induction Alternating Current (AC) motor drive system. Hybrid power supply systems are also considered. This chapter will give an electrical engineering background behind the railway vehicle modelling demonstrated in Chapter 4.
- Chapter 3 provides a review of relevant optimisation techniques. Numerical optimisation algorithms are considered in two main areas: unconstrained and constrained. Dynamic programming is also considered in this chapter due to its suitability for multi-stage optimisation. The optimisation methods discussed in this chapter are applied in each case study in Chapter 5, Chapter 6 and Chapter 7.
- Chapter 4 details the modelling and simulation of the traction power system. This chapter first discusses the physical equations determining vehicle motion. The concept of a vehicle state switch is then introduced. Coupled with train motion modelling, an energy consumption model is developed. A single train simulator developed in this chapter will be used as a basic tool for the case studies in following three Chapters.
- Chapter 5 introduces the first case study, that of the optimisation of a single train speed trajectory. This chapter first considers the optimisation of train speed trajectories from an optimal-control viewpoint. A train speed trajectory search model is presented, to which various algorithms are applied to search for energy efficient speed control strategies. The results gained in this chapter can be used as the fundamental information for a multiple train simulation in Chapter 6 and a Diesel Multiple Unit study in Chapter 7.



#### 1.4. THESIS STRUCTURE

---

- Chapter 6 covers the second case study on improvements to system energy efficiency for multiple trains on an electric network. A DC electric railway network is modelled using the Load Flow (LF) method. The GA is applied to search for the best operational strategy.
- Chapter 7 discusses the third case study on power management strategies for a DMU train. Dynamic Programming and an online rule-based method are proposed to search the optimal power management strategies.
- Finally, conclusions on major contributions and further work of the research demonstrated in this thesis are summarised in Chapter 8.

# Chapter 2

## A review of railway traction systems

### 2.1 Introduction

Like most road running vehicles, a railway traction power system provides mechanical power that can be converted into kinetic and potential energy and can be used to overcome resistance to motion [51].

For every successful train traction system, there are certain general requirements which should be met:

1. The capability to start and haul a prescribed load over its specified routes whilst maintaining the timetable;
2. A sufficiently long service lifetime with minimal non-availability due to maintenance and standby;
3. To be fuel efficient;
4. To be environmentally friendly.

Three typical railway traction systems are introduced in this chapter: electric, diesel electric and hybrid. DC and AC electric traction systems are commonly deployed in the railway industry. Very different power transmission system requirements exist for each system. Multiple power electronic control stages are

commonly used for such traction systems in order to create a smooth power supply from the power network. Diesel electric traction systems are mainly found on routes that are not electrified. Hybridisation of conventional power systems using energy storage devices will reduce energy consumption. Such hybrid traction systems are currently still in the development phase [52].

## **2.2 Railway traction systems**

### **2.2.1 Electric traction**

The first practical application of electric traction dates back to the second half of the 19th century. At the early stage of development, DC motors, together with low-voltage DC traction lines, were the main traction power supply methods, due to their simple torque control characteristics. Subsequently, low-voltage DC transmission networks and high-voltage, low frequency ( $16\frac{2}{3}$  Hz and 25 Hz) AC transmission networks became the two major electric supply methods for traction power. The reason for the emergence of the AC transmission network was due to the inherent tractive characteristics of induction motors and the difficulties of supplying electrical power from a DC or single phase AC transmission line [53]. It was not until the 1950s, with the advancement of power electronics, that high voltage and industry frequency AC transmission became a reality. Since then, 25 kV single-phase at 50 or 60 Hz often replaces the 1.5 kV DC networks and 3 kV networks [54].

DC and AC electric trains need to be supplied with power through a current collection system. A traction power network is employed to supply the electrical power to the entire electrified rail network. There are generally two types of power supply system: DC and AC [55, 56].

For a DC power supply system, electrical power is usually supplied through a conductor rail which is adjacent to the running rails. The original advantage of a DC power supply system was because it is easy to control the vehicle-mounted

traction equipments. With the development of power electronics technology, DC supply is regarded as advantageous due to its compact size and weight of control devices. Typical voltage for a DC power supply system is between 600V and 1.5kV. This implies higher current going through the supply circuits and higher electrical loss. DC power supply system is mostly applied for urban and regional lines. In Chapter 6, to study the energy efficiency of the operations of multiple trains, a part of regional railway network is studied.

An AC power supply system is usually using an overhead line consists of a contact line carrying live current and catenaries, which is an insulated suspension system supporting the contact line. An overhead line is located at a certain height above the rails. Higher voltage level in an AC power supply system reduce the current and electrical loss. Also, fewer substations are required compared with the lower voltage in DC traction networks. Generally, it is economic to use such system for a high-speed or heavy-haul railways.

### 2.2.2 Diesel-electric traction

The diesel engine was introduced in the early 20th century by Rudolph Diesel. Conventional diesel engines are not able to operate well at low speed, therefore a direct connection between the engine and the driving wheel is usually undesirable. Nowadays, power is usually transmitted to the wheels through a form of electric transmission, due to the convenience of operation [55]. In a diesel-electric traction system, the diesel engine is used to drive an electric generator, which then conveys the electrical power to the traction motors. An architecture diagram for a typical diesel-electric propulsion system is shown in Figure 2.1.

Compared to the electric propulsion system, diesel electric propulsion is less powerful and less efficient with high maintenance costs, more pollution and noise. However, diesel electric propulsion is used for a large range of applications in the railway transportation industry. Only 33% of railway routes are electrified in the UK; diesel electric traction therefore plays a key role in the British railway

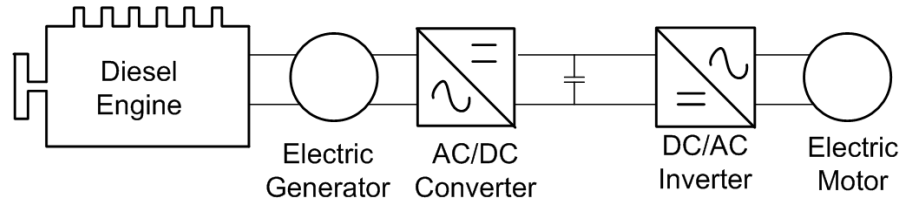


Figure 2.1: An architecture diagram for a diesel-electric propulsion system

industry [57]. Some of the advantages of diesel electric traction are summarised as follows [49, 58]:

1. The elimination of the dependency on an electrical power supply, this means that trains are not restricted to certain routes due to the installed power supply;
2. The removal of failure modes and maintenance activities associated with the power supply and electric traction system;
3. No capital investment is required for a fixed power system; Although diesel trains are more expensive, on low usage lines, a clear business case can be made for diesel as compared to electric traction;
4. A diesel engine obtains its highest efficiency at a particular range of speeds; It is feasible to reduce the fuel cost by applying power management strategies.

### 2.2.3 Hybrid traction

Hybrid railway vehicles can be defined as follows: railway vehicles, i.e. locomotives, or multiple unit trains, which use a fueled power system for propulsion and are supported by an on-board rechargeable energy storage system. The fuel cell hybrid vehicle [41, 42] and the more conventional diesel-electric hybrid [59] are two of the most common forms of hybrid railway vehicle.

The following two key advantages of on-board energy conversion technology could be used to reduce overall energy consumption and gas emission [60, 61]:

- Utilisation of regenerative energy. A rechargeable energy storage system will effectively store the regenerative energy during braking to be used later for traction.
- Improvement of efficiency for the operation of the prime mover. For example, the engine could be switched off if it is running with low energy efficiency and the energy storage system could provide the traction power.

On-board and wayside energy storage has been attracting an increasing amount of interest for the application of the hybrid railway vehicle. Sone states in his report that two problems need to be solved through on board or wayside energy storage devices [62]. These two problems have existed since AC-motored Electric Multiple Units (EMUs) with regenerative braking capability were introduced. Energy storage devices can be used to achieve the following three goals during regenerative braking:

- Manage over-voltage at the traction motor without significant increase in pantograph voltage;
- Reduce over-current from the pantograph to the feeding system;
- Improve the usage rate of regenerative power.

An architecture diagram for a diesel hybrid propulsion system is shown in Figure 2.2. A diesel engine can be replaced by another type of prime mover, such as fuel cells [63].

## 2.3 DC motor drive

The most common DC traction motors used in the railway industry are series and separately-excited machines.

In the series motor, the field winding is connected in series with the armature, resulting in the field being determined by the armature current. The series

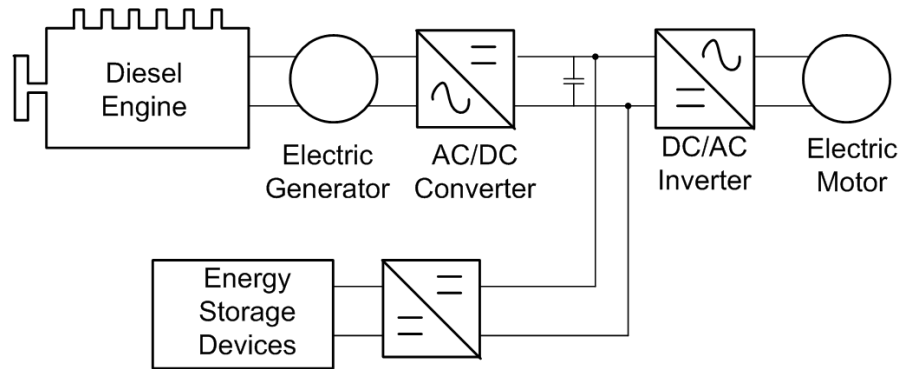


Figure 2.2: An architecture diagram for a diesel hybrid propulsion system.

field forms a protection to ensure that no current flows in the motor without a corresponding field having already been established. The field becomes non-linear when iron saturation occurs at a higher current. For any series traction motor with a steady terminal voltage, “wheel-slip correction” can be inherently realised with a high starting torque followed by a constant power operation regime with increasing speed. This is a very attractive feature for stable operation. Speed control for a series DC machine can be achieved by varying the terminal voltage at the starting stage and the field through the diverter resistor at a greater speed which is usually referred to as a “field weakening operation”. However, the insertion of a series resistor into the motor circuit is regarded as wasteful of power and of limited use.

In a camshaft controlled train, the series resistance is varied to give the overall characteristic shown in Figure 2.3. At the beginning, both armature and field current are both maintained constant to generate a constant torque. With the increase of rotating speed, the output power is increased until the base speed. Subsequently, field current is reduced and bring down the output torque. This operation called field weakening operation. After the rotating speed is over the transition speed, the armature current begins to drop and further bring down the torque. The motor then goes into a weak field operation.

For the separately-excited motor, the field excitation circuit is independent from the field circuit. Speed control is also realised through varying the armature voltage

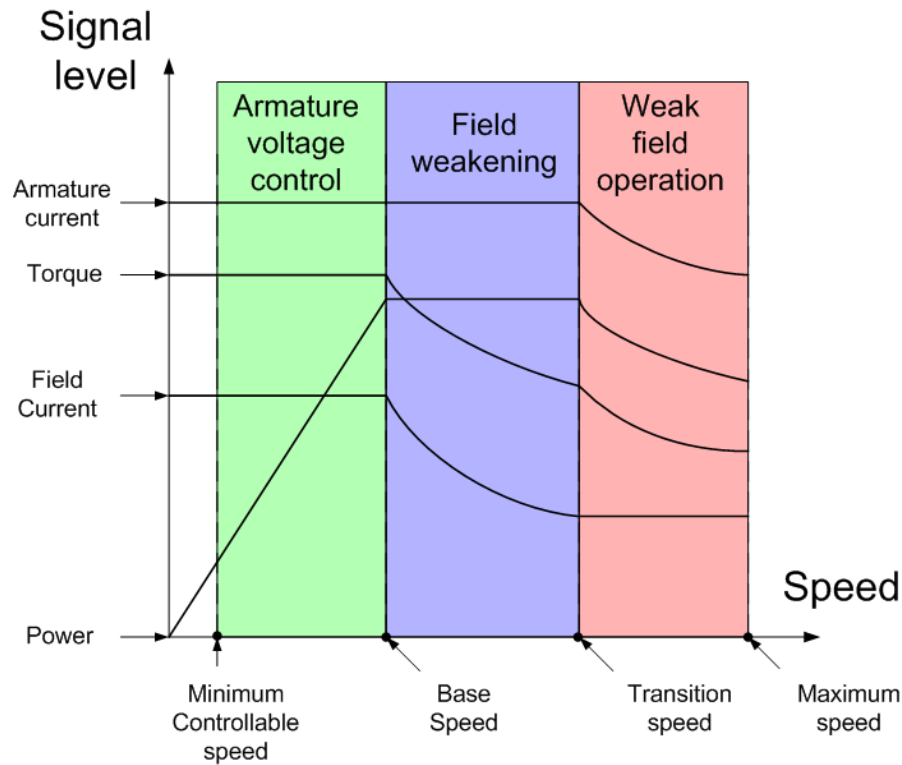


Figure 2.3: Separately excited DC motor traction drive regime of operation

or the field current. This type of motor with its separately excited field can be used for regenerative braking. In Figure 2.3, three control regimes have been identified with different characteristics in terms of torque, current, speed and power [54].

Further details of DC motor drive can be found in Appendix A.1. A.1

## 2.4 AC motor drive

The induction motor has long been regarded as a suitable final drive for railway traction systems due to its inherent capability for regeneration and steep speed and torque characteristics. However, the widespread introduction of induction motors could only be realised after modern power electronics became available [53]. The two reasons for this are as follows:

- Speed control of the induction motor is achieved by changing the input frequency and voltage of the input power supply;



- The difficulty of obtaining the proper three-phase supply from a DC or single-phase AC traction line should be solved by power electronics techniques.

The development of power electronics relies on advancement of semiconductor devices [53]. In the 1960s, the development of the power thyristor gave rise to trials of several experimental inverter-fed induction motor locomotive drives. The main drawback of this type of device is the lower current and voltage level which it could withstand which limits its application in high power fields. In the 1970s, the development of a high-power force-commutated thyristor led to the deployment of the Current Source Inverter (CSI) in DC-fed metro traction drives. Later, the large power Gate-Turn-Off (GTO) thyristor realised the Voltage Source Inverter (VSI) in railway applications. Until very recently, the Insulated Gate Bi-polar Transistor (IGBT), which has a lower operation current and higher switching speed has taken the place of the GTO. In the 1980s, the pulse converter was developed, which is a four-quadrant AC-DC line converter that enables three-phase VSI-Induction Motor drives on single-phase AC supplied railways.

There are two commonly applied types of AC motor: synchronous and asynchronous.

In synchronous motors, the armature winding is on the stator, inducing the main voltage, and the field winding is on the rotor, producing the main magnetic field. Such roles of the winding are converse to normal practice. The symmetric three-phase sinusoidal currents in an armature winding produce a uniform rotating magnetic field. The DC current through the field winding produces a steady-state magnetic field. The rotor field tries to line up with the stator field which produces torque. The larger the angle between the two magnetic fields, the greater the torque on the rotor [64]. The synchronous motor was once used in electrical railway traction and was replaced by the second type of AC motor, until GTOs were introduced [53].

The asynchronous motor is also referred to as an induction motor since the rotor voltage and rotor magnetic field are induced rather than being connected physically

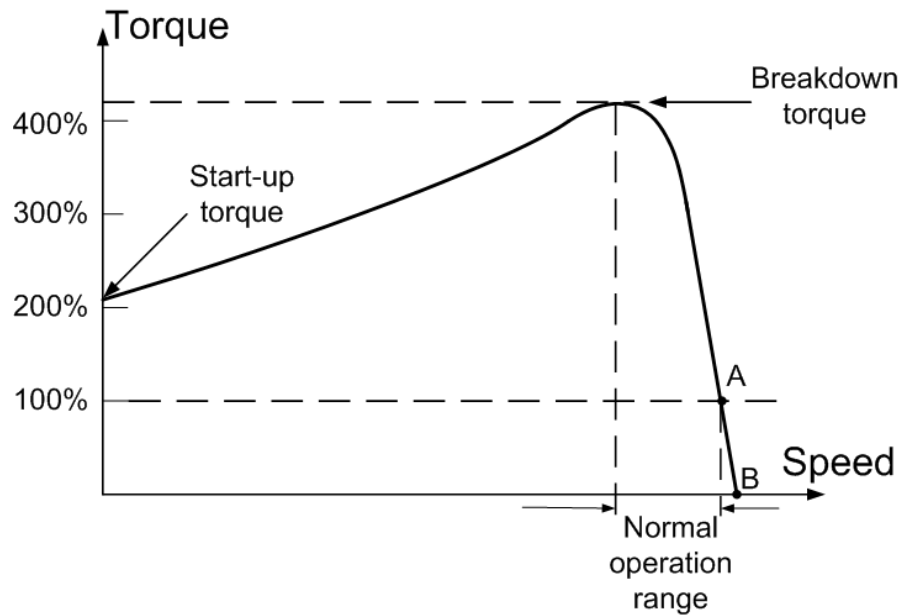


Figure 2.4: Typical induction motor torque-speed characteristic curve

by wires [64]. The significant difference between asynchronous and synchronous motors is that the DC current is not required in the rotor winding. The three-phase current through stator induces a rotating flux in the air-gap. Current is then induced within the rotor conductor, setting up an opposite air-gap flux to this rotating flux vector. The torque is produced due to the reaction between the rotor current and net air-gap flux.

The following comments are made according to the individual induction motor torque-speed characteristic curve [64], as shown in Figure 2.4. All the torque shown in Figure 2.4 are represented by percentage of full-load torque which is the torque of point A.

1. The torque of the motor is zero at synchronous speed which is the speed for point B.
2. The torque-speed curve is nearly linear between no load (Point B) and full load (Point A).
3. There is a maximum torque and it is referred to as “breakdown torque” or “pullout torque”.

4. The starting torque on the motor is slightly larger than its full-load torque. The motor is able to take up any load that it can supply at full power.
5. If the speed of motor goes higher than synchronous speed, the torque will switch its direction, implying that the machine switches to generator mode under regenerative conditions.

Control of the speed of the motor is achieved through applied voltage and stator frequency, both of which are necessary. By using variable frequency control, it is possible to change the speed of the motor to either above or below the base speed shown in 2.5. It is necessary to reduce the terminal voltage applied to the stator to avoid magnetic saturation in the core of the induction motor and excessive magnetisation currents flowing through the machine. This process is referred to as *derating* [53]. Vector control is one of the methods used in variable-frequency induction motor drives to control the torque and hence the speed. In this control method the stator current vector is transformed into rotor flux linkage current, i.e. excitation current, vector and rotor torque current vector [65].

Figure 2.5 shows different control regimes of operation. Firstly, it is necessary to introduce the relationship between flux and the voltage-to-frequency ratio  $V_a/\omega$ ,

$$V_a(t) = E_a \sin(\omega t) = -N_p \frac{d(\phi)}{dt} \quad (2.1)$$

$$\phi(t) = \frac{E_a}{\omega N_p} \cos(\omega t) \quad (2.2)$$

where  $E_a$  is the air-gap Electromotive Force (EMF), which is approximately the same as terminal voltage  $V_a$  except at low speed when more voltage drop is introduced between  $E_a$  and  $V_a$ .  $\phi$  is the magnetic flux and  $N_p$  is a constant.

As stated in (2.2), flux is proportional to the voltage-to-frequency ratio. A constant flux should be maintained for the start-up stage and constant flux region for a constant torque operation and can be achieved by keeping voltage-to-frequency ratio constant.

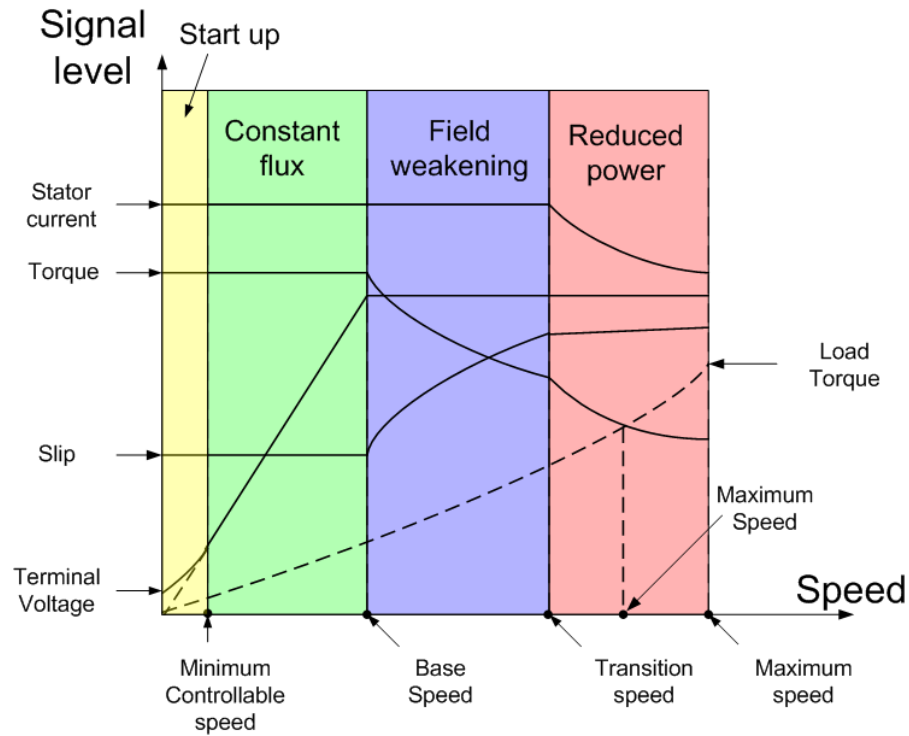


Figure 2.5: Mechanical and electrical variables over traction duty cycle

From low to higher speed, various operation regions are identified throughout the whole traction duty cycle.

- **Start up region:** The stator impedance voltage drop increases when the stator resistance becomes comparable to the reactance at low speed. As a result, the air-gap flux is reduced and  $V_a/\omega$  must be increased to maintain a constant voltage-to-frequency ratio.
- **Constant flux region:** For a given current through the rotor, the torque in the machine is proportional to the air-gap flux density. Consequently, the constant flux density results in constant torque operation.
- **Field weakening region:** If the speed increases further above the base speed, the frequency will increase further whereas the terminal voltage remains unchanged. This process is termed as field weakening, as both the air-gap flux and available torque reduce during this period. Both the current and voltage are maintained as constant while machine slip increases. This stage will last

until the working torque equals breakdown torque as shown at 2.4.

- **Reduced power region:** At this stage further speed increases are obtained by increasing the stator frequency while the line current decreases. Power consumption will reduce and the slip will increase by a small amount until the balancing speed is achieved at which the maximum output torque equals to load torque.

Details of different commonly seen induction motor drive can be found in Appendix A.2

## 2.5 Summary

This chapter has reviewed railway traction systems, including the power supply and final DC and AC drives and diesel-electric systems. Furthermore, with significant potential in the application to the hybrid railway vehicle, both of electric and hybrid supplies remain interesting if rechargeable Energy Storage Devices (ESDs) are included. Separate excitation and series excitation are the two most popular types of excitation in the DC motor drive. In particular, series excitation has inherent advantages in magnetic field protection and wheel slip correction.

Two power electronic control systems for the DC motor drive are introduced. There are two types of AC motor: synchronous and asynchronous. The latter type is also referred to as the induction motor, which is more popularly applied to railway traction. The duty cycle of operation of induction motors is introduced.

Multiple power systems, such as engines and motors, are commonly adopted on modern railway vehicles. The operations between these sources are usually considered to be homogenous, however, distinctive operation can be possible if the sources can be decoupled and work independently. Intuitively, power outputs from all sources are summed up to meet the power demand from the driver or the supervisory control unit. For a certain power demand, rather than making an equal

## 2.5. SUMMARY

---

division between the power systems as usual, a different division is made, which gives rise to another problem: “how is the division of power demand made?”. Such a problem can only be answered with further study of the characteristics of the power system and various vehicle running duty cycles.

# Chapter 3

## Review of optimisation techniques

### 3.1 Introduction

Optimisation generally refers to the procedure of selecting one feasible solution to minimise or maximise the objective function [66, 67]. Let the objective function “ $J$ ” be a real valued function of “ $n$ ” real variables  $x_1, x_2, \dots, x_n$ . The basic form is shown in Eqn. 3.1:

$$f(x) \rightarrow \text{Extreamum} \quad \text{Subject to} \quad \begin{aligned} g_i(x) &\geq b_i \quad i \in \mathcal{I} \\ h_j(x) &= b_j \quad j \in \mathcal{E} \end{aligned} \quad (3.1)$$

where  $f(\cdot)$  is the function to be optimised,  $g_i(\cdot) \ i = 1, 2, \dots, m$  is the  $m$  functions of the inequality constraints <sup>1</sup> and  $h_j(\cdot) \ j = 1, 2, \dots, k$  are the  $k$  constraint functions of equality.  $\mathcal{I}$  and  $\mathcal{E}$  are sets of indices for equality and inequality constraints.

If the objective function  $f(\cdot)$  is to be minimised,  $f(\cdot)$  is also referred to as the “cost function”; if it is to be maximised,  $f(\cdot)$  is known as the fitness function [68]. In this chapter,  $f(\cdot)$  is usually considered as a “cost function” to be minimised.

For example, as mentioned in Chapter 2, how the division of power demand is made will affect the total fuel consumption, even if the entire output of power is the same. In this case, total fuel consumption is regarded as an objective function to be

---

<sup>1</sup>inequality constraints can also be “ $>$ ” rather than “ $\geq$ ”

minimised, while engineering constraints should be imposed to act as constraints. The study tries to identify the energy demand distribution between engines, i.e. “x” in 3.1, to minimise the total fuel cost.

The optimisation algorithms can generally be categorised into two main sub-fields. One is numerical optimisation algorithms shown in 3.2 and the other is metaheuristics introduced in 3.4. In addition, dynamic programming is introduced in Section 3.3 due to its importance in the multiple decision making applications.

### 3.2 Numerical Optimisation

The characteristics of numerical optimisation algorithms are summarised as follows:

- Numerical optimisation algorithms are heavily affected by the characteristic of the objective function and the constraints;
- The derivative information of the objective function is required for most numerical optimisation algorithms;
- The searching strategies of the numerical algorithms are deterministic, which means that the consequences of the searching action are predictable [69];
- The numerical optimisation algorithms usually take one of the feasible solutions as the initial solution at the beginning iterative procedure; as a result, only sub-optimal results can be found for a non-convex objective function.

For a numerical optimisation problem, a starting point  $x_0$  is required. The selection of  $x_0$  will significantly affect the performance of the algorithms. Algorithms generate a set of solution  $\{x_i\}_{i=0}^{\infty}$ . Algorithms will be terminated if no decrease is achieved by the generated solution or if  $\|f(x_i) - f(x_{i+1})\| \leq \epsilon$ , where  $\epsilon$  is sufficiently small, i.e. smaller than a arbitrarily selected small value.



The other main function of algorithms is to decide how to generate  $x_{i+1}$  from  $x_i$  for  $i = 0, 1, \dots$ . The information used is obtained from the objective function  $J$  at  $x_i$  or also the information from earlier iterations  $x_0, x_1, \dots, x_{i-1}$ . The new solution produced should be finally able to decrease the value of the objective function in a finite number of iterations which are  $f(x_i) < f(x_{i-k})$  for  $k = 1, 2, \dots, i-1$ . There are two commonly used strategies to handle the selection of the next candidate point: line search and the trust region [66].

In line search strategy, the generation of each point  $x_i$  depends on two pieces of information: the search direction  $d_i$  and the step length  $\pi_i$ . The iteration is shown in Eqn. 3.2.

$$x_{i+1} = x_i + d_i \pi_i \quad (3.2)$$

where  $\pi_i \geq 0$ . Obviously, the direction  $d_i$  is required to be a descent direction for function  $f(\cdot)$  to ensure the iteration will lead to the final reduction of the function value.

Take a typical descent direction so that  $d_i^T \nabla f_i < 0$ . Assume that  $d_i$  takes the general form as follows:

$$d_i = -M_i^{-1} \nabla f_i \quad (3.3)$$

where  $M_i$  is symmetric and positive definite matrix,  $\nabla$  is the first-order partial derivatives of  $f_i$ . Generally, according to the different selection of  $d_i$  the line search will be further categorised into following three types.

**Steepest descent method [70]** This method is also called “gradient descent method”.

In this method,  $M_i$  is the identity matrix  $I$  so that the searching is along the steepest descent direction.

$$d_i = -\nabla f_i \quad (3.4)$$

$$x_{i+1} = x_i - \nabla f_i \pi_i \quad (3.5)$$

**Newton’s method**  $M_i$  is the Hessian  $\nabla^2 f_i(x_i)$ . The derivative is zero at a minimum

or maximum. According to the Taylor's expansion of  $\nabla f(x)$  in Eqn. 3.6.

$$\nabla f(x + \Delta x) = \nabla f(x) + \nabla^2 f(x) \Delta x \quad (3.6)$$

The extremum is obtained by setting Eqn. 3.6 to zero. Note that the necessary condition of  $x^*$  being a local minimiser of  $f$  and  $\nabla^2 f(x)$  being continuous and existing is:  $\nabla f(x^*) = 0$  and  $\nabla^2 f(x^*)$  is semi-positive definite. So the iteration becomes:

$$d_i = -\frac{\nabla f_i(x_i)}{\nabla^2 f_i(x_i)} \quad (3.7)$$

$$x_{i+1} = x_i - \frac{\nabla f_i(x_i)}{\nabla^2 f_i(x_i)} \pi_i \quad (3.8)$$

**Quasi-newton method**  $M_i$  is the approximation of the Hessian  $\nabla^2 f_i(x_i)$  and it is to be updated for every iteration. The Taylor series of  $f(x)$  is approximated as:

$$f(x_i + \Delta x) \approx f(x_i) + \nabla f(x_i)^T \Delta x + \frac{1}{2} \Delta x^T M_i \Delta x \quad (3.9)$$

The gradient of  $f(\cdot)$  can be approximated by

$$\nabla f(x + \Delta x) \approx \nabla f(x) + \nabla^2 f(x) \Delta x \quad (3.10)$$

By setting this approximation of gradient to zero, we have:

$$d_i = -\frac{\nabla f_i}{M_i} \quad (3.11)$$

$$x_{i+1} = x_i - \frac{\nabla f_i}{M_i} \pi_i \quad (3.12)$$

Various methods are used to find the approximated Hessian  $M_i$  which is symmetric.

The description of the above three line search methods [66, p.19] focuses on the selection of direction  $d_i$  at each iteration. Substituting Eqn. 3.3 into  $d_i^T \nabla f_k$ , gives:

$$d_i^T \nabla f_i = -\nabla f_i^T M_i^{-1} \nabla f_i < 0 \quad (3.13)$$

Note that  $M_i$  is symmetric and positive definite. In this case, at each iteration  $d_i$  is the descent direction. The step length  $\pi_i$  can be calculated by approximately obtaining the minimum of the following formula.

$$f(x_i + d_i \pi_i) \quad (3.14)$$

To maintain the descent direction,  $\pi_i$  is set to be positive. In practice, the calculation of the real minimum of Eqn. 3.14 is computationally expensive and unnecessary. The step length is usually determined after certain conditions, i.e. the Wolfe Conditions, are met [66]. A so-called backtracking approach is also used to determine the step length.

Another searching strategy is “trust region” [66, p.19]. In this strategy, a model function  $f'_i(\cdot)$  is constructed to approximate the behavior of the objective function  $f(\cdot)$  near the current iterative point  $x_i$ . In order to restrict the approximation in a small area around function  $f(\cdot)$ , a small region around  $x_i$  is defined. This small region is called the “trust region”.

The model function  $f'_i(\cdot)$  is usually a quadratic function as follows:

$$f'_i(x_i + \pi) = f_i + \pi^T \nabla f_i + \frac{1}{2} \pi^T M_i \pi \quad (3.15)$$

where the  $f_i(\cdot)$  and  $\nabla f_i(\cdot)$  are the objective function and its gradient respectively and  $M_i$  can be the Hessian or approximation of Hessian for function  $f_i$ . Eqn. 3.15 is a Taylor Series of  $f_i$  while deleting the elements with a higher derivative order than 2. After the model function and “trust region” are both defined, a candidate step  $\pi_i$

is obtained through solving the following formula:

$$\min_{\pi_i} f'_i(x_i + \pi_i) \quad (3.16)$$

Note that the  $x_i + \pi_i$  is within the predefined trust region. A trust region which is too large may result in an insufficient decrease of the objective function  $f(\cdot)$ . Reduction of the region in the next iteration is the consequence. The trust region can be defined by Eqn. 3.17

$$\|\pi_i\|_2 \leq \Delta R_i \quad (3.17)$$

where  $\Delta R_i \geq 0$  is referred to as the trust-region radius [66, p.19].

In summary, both strategies are distinguished from each other in the selection of direction and step length. Line search determines the direction  $d_i$  first and then solves Eqn. 3.14 to identify the step length  $\pi_i$ . In the case of the trust region strategy, the direction does not follow a fixed pattern and is determined after the step length is identified. The minimisation procedure shown in Eqn. 3.16 will first try the maximum step length, which is the trust-region radius  $\Delta R_i$  and look for the step length and direction set with the most significant reduction of  $f(\cdot)$ . If searching is not satisfactory,  $\Delta R_i$  will be further reduced. Categories of a group of numerical optimisation methods can be found in Appendix B.

## 3.3 Dynamic programming

### 3.3.1 Introduction

Dynamic Programming (DP) was firstly proposed and later developed by American mathematician Richard Bellman in the 1950s [71]. Dynamic programming is an appropriate method for discrete multi-stage decision optimisation problems. For each decision, the sub-problem can be solved in a similar fashion. An optimal solution of the original problem can be found from optimal solutions of every sub-

problem [72].

The dynamic programming method is based on Bellman's Principle of Optimality [73] defined in Def. 3.1.

**Definition 3.1** (Principle of Optimality). *An optimal policy has the property that whatever the initial state and the initial conditions are, the remaining decisions must constitute an optimal policy with regards to the state resulting from the first decision.*

This assertion states that optimal policies would be derived from the suboptimal policies, whose improvement will give rise to the improvement of whole policies.

This section firstly presents the mathematical formulation of dynamic programming, the two key characteristics of DP are then discussed and the general steps to solve a problem using DP is summarised.

#### 3.3.2 Mathematical presentation

##### Objective function

For the dynamic programming method, the objective function for “n” decision stages has the following forms.

$$h(d_1, d_2, \dots, d_n) = C_1(d_1) \circ C_2(d_2) \circ \dots \circ C_n(d_n) \quad (3.18)$$

Here, the sub-cost function  $C_n$  is defined as:

$$C_i = f_c(s_i) \circ f_i(s_i, d_i) \quad (3.19)$$

where:

- $s_i$  is the system state at  $i$ th stage and it is determined by the decision set  $\{d_1, d_2, \dots, d_{i-1}\}$ . In particular, the initial state  $s_1$  is pre-determined;
- $d_i$  is the decision made at  $i$ th stage;

- $f_c$  is the system cost function due to  $s_i$ ;
- $f_t$  is the system transition function due to  $s_i$  and  $d_i$ ;
- $\circ$  is the associative binary operation, e.g. additive or multiplicative operation.

### Sequential decision process

Many optimisation problems require a set of decisions  $\{d_1, d_2, \dots, d_n\}$  to be found to minimise the objective function  $h$ . Let  $\Delta_n$  denote the feasible decision space at the  $n$ th stage,  $D^*$  denote the optimum decision set  $\{d_1^*, d_2^*, \dots, d_n^*\}$  at which the objective function in Eqn. 3.18 achieves its minimum and  $H^*$  denote the minimum objective function output.

$$H^* = h(D^*) = h(d_1^*, d_2^*, \dots, d_n^*) \longrightarrow \text{minimum} \quad (3.20)$$

An intuitive way to find the  $D^*$  is the “brute force” method. The major philosophy behind this method is to try out all the possible combinations of feasible decisions at each stage and select the minimum cost caused by every combination. This is not impossible for very small problems, however, it can cause an unnecessary computational burden. Also for some complex problems it may not give an answer in an acceptable time scale.

Another way to tackle this type of problem is to divide the problem into sub-problems. The aim is then to solve the sub-problems by optimising the sub-objective function first and then achieve the optimum for the entire problem. For a multi-stage decision making problem, dividing the problem by the stage is an intuitive approach.

$$\begin{aligned} H^* &= \min_{(d_1, d_2, d_3, \dots, d_n) \in \Delta} \{h(d_1, d_2, d_3, \dots, d_n)\} \\ &= \min_{d_1 \in \Delta_1} \{\min_{d_2 \in \Delta_2} \{\dots \{\min_{d_n \in \Delta_n} \{h(d_1, d_2, \dots, d_n)\}\} \dots \}\} \end{aligned} \quad (3.21)$$

where  $\Delta = \Delta_1 \times \Delta_2 \times \dots \times \Delta_n$  and  $d_i \in \Delta_i$  Eqn. 3.21 are referred to as “*sequential*”

### 3.3. DYNAMIC PROGRAMMING

---

decision processes” [72]. Note that each decision made depends on the decision made for the previous stages, i.e.

$$\Delta_i \stackrel{\text{def}}{=} f_{\Delta}(\Delta_1, \Delta_2, \dots, \Delta_{i-1}) \quad (3.22)$$

In Eqn. 3.22 it is stated that the decision space is defined by the decision in the past. If  $\Delta_i$  is only defined by  $\Delta_{i-1}$ , this decision process is also called the “*Markov Decision Process*” [74, 75].

Each of the “min” functions can be interpreted as a procedure to seek an optimum of the sub-problem.

$$\begin{aligned} & \min_{d_i \in \Delta_i} \{h(d_1, d_2, \dots, d_{i-1}, d_i, d_{i+1}, \dots, d_n)\} \\ & = \min_{d_i \in \Delta_i} \{h(d_1, d_2, \dots, d_{i-1}, d_i, d_{i+1}^*(d_i) \dots, d_n^*(d_i))\} \end{aligned} \quad (3.23)$$

where  $d_{i+1}^*(d_i)$  is the optimum decision resulting from the decision made at the “ith” stage. The equation shown in Eqn. 3.23 is the minimisation procedure at the “ith” stage to find an optimum decision policy that will minimise the cost due to the decision set  $\{d_i, d_{i+1}, \dots, d_n\}$ . A function to describe the resulting cost by the following decisions to the current decision  $d_i$  is referred to as the cost-to-go or value function [76, p.12]. Let function  $f(s_i)$  denote the solution of the cost-to-go function where  $d_i, d_{i+1}, \dots, d_n$  are to be made, and correspondingly, let  $f^*(s_i)$  denote the optimum solution. We have:

$$f^*(s_i) = \min_{\Delta_i} \{ \min_{\Delta_{i+1}} \{ \dots \{ \min_{d_n} \{ C_i(s_i, d_i) \circ C_{i+1}(s_{i+1}, d_{i+1}) \circ \dots \circ C_n(s_n, d_n) \} \} \dots \} \} \quad (3.24)$$

Now substituting Eqn. 3.18 and into Eqn. 3.21, it is deduced that:

$$\begin{aligned}
 H^* &= \min_{d_1 \in \Delta_1} \{ \min_{d_2 \in \Delta_2} \{ \cdots \{ \min_{d_n \in \Delta_n} \{ h(d_1, d_2, \cdots, d_n) \} \} \cdots \} \} \\
 &= \min_{d_1 \in \Delta_1} \{ C_1(d_1, s_1) \circ \min_{d_2 \in \Delta_2} \{ C_2(d_2, s_2) \circ \cdots \circ \min_{d_n \in \Delta_n} \{ C_n(d_n, s_n) \} \} \} \\
 &= \min_{d_1 \in \Delta_1} \{ C_1(d_1, s_1) \circ f^*(s_2) \}
 \end{aligned} \tag{3.25}$$

Eqn. 3.25 can be solved recursively since  $H^* = f^*(s_1)$ .

### 3.3.3 Elements of dynamic programming

It is summarised in [77] that two essential elements must be found in the problems to which DP becomes applicable: optimal substructure and overlapping sub-problems.

#### Optimal substructure

Optimal substructure can be defined as follows.

**Definition 3.2** (Optimal substructure). *A problem demonstrates optimal substructure if the optimal solution to this problem contains the optimum solutions to sub-problems.*

In Eqn. 3.25, the optimum solution to the problem contains within it an optimum solution  $H^*$  to a sub-problem which is  $f^*$ . To identify the characteristics of “Optimal substructure” defined in 3.2, a common pattern is suggested [77]:

1. A solution to a problem can be demonstrated as a set of decisions;
2. A decision made can be regarded as a choice possibly leading to an optimal solution; it is not yet known whether this decision will or not do so;
3. It can be proved that the solutions to the sub-problems in the optimal solution to the problem should be optimal by themselves through the contradiction strategies.



Take the shortest path problem for an example where a set of points along the path should be selected. Each path as a solution consists of all the possible points between the start point to the end point. Selecting which point to go for each step can be regarded as a decision to make. But whether it is an optimal solution is unknown until all the decisions have been made. For an optimal path, each subset of points along the route should be also be the shortest path. If not, the subset of points can be swapped with the optimal ones which will contribute to a shorter path of whole route. This will contradict the original assertion saying it is an optimal solution.

#### **Overlapping sub-problems**

Overlapping sub-problems are the other key ingredient for DP. This character implies that each of the sub-problems should be kept in a similar format so that a recursive algorithm can be applied.

**Definition 3.3** (Overlapping sub-problems). *In the procedure of solving a problem, if there exists an algorithm which is able to solve the same sub-problem repeatedly, it is claimed that this problem has overlapping sub-problems.*

If a problem has overlapping sub-problems, it is possible to apply the same algorithm repeatedly. DP solves each sub-problem once and stores the optimal solution in a look-up table. The look-up table can be retrieved in a approximately constant time and this will reduce the computational time for the recursive algorithm. The storing process is referred to as a “memoisation” [77]. A memoised recursive algorithm maintains an entry for the solution to each sub-problem. Initialisation is needed to indicate the occurrence of storing a new solution. The memoisation procedure for one sub-problem is terminated until the optimum solution is found. In the subsequent step, the value stored in the look-up table can be simply referred to.

#### 3.3.4 Dynamic programming and the greedy algorithm

In dynamic programming, each choice is made based on the past solution for the sub-problems. This bottom-up problem solving style leads DP to progress from smaller sub-problems to the larger problem. In a greedy algorithm, the best choice which is available at the moment is taken. After the choice has been made, the problem will progress into a smaller problem and a new choice will be made depending on the choices made so far, but not on the solutions of sub-problems. In that sense, a greedy algorithm progresses in a top-down fashion. A greedy algorithm can only achieve the optimum solution if it can be shown that each step yields a globally optimum solution. This is referred to as the “greedy-choice property” [77] and it only exists in a range of situations, e.g. shortest path searching using Dijkstra’s algorithm [78]. To summarise, to apply the greedy algorithm, according to [77, p.380] the following steps should be followed:

- The optimisation problem can be modelled as one for which the choice can be made with one sub-problem left to be solved;
- Prove that the greedy choice is logical because there is an optimal solution to the original problem by making the greedy choice; this step can be regarded as the proof for the greedy-choice property;
- After a greedy choice is made, a sub-problem will be left; if we assume that there is an optimal solution to the sub-problem, the combination of the greedy choice and the optimal solution will be the optimal solution to the original problem.

#### 3.3.5 Summary

As DP solves the problem in a bottom-up fashion as opposed to the greedy algorithm, simpler or smaller problems will be solved before it is used by other, more complex and bigger problems [79]. To summarise, the DP algorithm consists of 4

steps which are as follows [80]:

1. A recursive definition for the optimal solution as in Eqn. 3.25;
2. A look-up table to store the optimal solution for the sub optimal problem;
3. A bottom-up approach which starts from the simplest sub-problem to fill the look-up table;
4. A track-back method to reconstruct the final optimal solution to the problem.

## 3.4 Metaheuristics

### 3.4.1 Introduction

In the applications where numerical optimisation is not suitable or not effective enough, e.g. combinatorial optimisation in which case the set  $X$  is discrete or even finite [81], heuristic information is usually used to direct the search for the optimum answer. These kinds of algorithms are generally regarded as metaheuristics. In computer science, many optimisation problems with certain practical significance are “NP-complete”, for which no polynomial-time algorithm has yet been discovered. As a result, the numerical algorithms are unable to achieve the optimum answer in a polynomial time. In [77], at least three approaches to tackle NP-completeness are introduced. One of them is to use the approximate algorithms to find near-optimal solutions. Most metaheuristics are regarded as approximate algorithms which can only achieve suboptimal solutions. However, computational time can be significantly reduced.

Metaheuristics, first mentioned in [82] are originally defined as the solution methods to create a process capable of escaping from local optima and performing a robust search of a solution space. The interaction between local improvement procedures and higher level strategies is considered. More generally, metaheuristics include any procedures which use strategies to overcome the local optima and

achieve a better quality solution in complex solution spaces. The commonly employed processes in metaheuristics are to define the admissible move from one solution to another at each iterative step or to take each step by creating or destroying a new solution [83, p.vi] [84, p.47].

There are a variety of successful metaheuristics, such as simulated annealing [85, 86], Tabu Search (TS) [82, 87–89], guided local search [90, 91], Greedy Randomised Adaptive Search Procedures (GRASP) [92, 93], genetic algorithms and other evolutionary algorithms [94], the ant colony algorithm [84], iterated local search [95], scatter search [96] etc.

One dimension of the classification of metaheuristics concerns the degree to which the neighbourhood is exploited during the searching procedure. In GA, the neighbourhood is implicitly defined and is to be exploited by exchanging the components of one solution with those of another. However, for the other population based methods, e.g. scatter search, path re-linking enables the neighbourhood structures to be used in full generality. For other metaheuristics manipulating only a single solution, e.g. tabu search and, simulated annealing, the neighbourhood is searched based on various forms of randomised or deterministic strategies [83, p.vi]. In [84], metaheuristics is categorised into two classes: constructive and local search based. The number of solutions that is manipulated can classify metaheuristics into two groups: single solution and multiple solutions. The performance of some metaheuristics such as Ant Colony Optimisation (ACO) and GA will be greatly enhanced if local search can be included.

In Section 3.1, it is stated that all metaheuristics are approximate algorithms. Metaheuristics is currently one of the effective ways of tackling a range of NP-complete problems which cannot be solved in a polynomial time [77, Cha.35]. Some of the other characteristics of metaheuristics are summarised as follows:

- Metaheuristics do not require the gradient or Hessian matrix of the objective function. As a result, the objective function does not need to be continuous or differentiable and it can have constraints.

- Since metaheuristics take different strategies to overcome the local optima, they are suitable to search for an optima for a non-convex objective function.
- There is no general termination criterion [84].

GA and ACO are introduced in Section 3.4.2 and Section 3.4.3 respectively. Some more metaheuristics such as Simulated Annealing and Tabu Search are also covered in Appendix B.

#### 3.4.2 Genetic algorithm

A GA employs some “operators” to model the natural evolution process. A GA maintains a set of feasible solutions to a problem and new solutions can be obtained by using the operators. The solutions should be selected according to their quality. Iteratively, until the termination criterion is met, A GA will keep generating a new set of feasible solutions in the hope that improvement in the solutions will occur [97].

A five-step procedure is usually adopted for a GA [98]. The first step to implement a GA involves the encoding of the objective function to present the “chromosome” on which the evolution process operates [99]. The second step is to define a fitness function or selection criteria. Because a GA operates with a population of possible solutions, the third step is to create a population of individuals [100]. The fourth step is the iterative evolutionary process which takes place during reproduction. In this step, a set of operators demonstrated in nature are adopted, including selection, crossover, mutation, inversion and fitness-proportionate reproduction. Further details of each of these strategies can be found in [100, 101]. The final step will be decoding the results to obtain the solution.

Before a pseudo-code for a typical genetic algorithm is given, some comments are listed as follows:

- Let  $POP$  denote the population which contains a set of individuals and let  $s_{bsf}$  denote the individual with the best fitness value so far;

- Function INITIALISE generates the initial population of individuals;
- Function EVALUATE works as the fitness function in the above-mentioned second step; it returns the fitness value of the each generated individuals;
- Function GET-BEST returns the individual with the best fitness value;
- Function REPRODUCE is a function which performs all the operators in the fourth steps and returns the new generation of individuals.

---

**Algorithm 3.1** Pseudo-code for Genetic Algorithm

---

```

POP ← INITIALISE()
EVALUATE(POP)
sbst ← GET-BEST(POP)
while Termination Requirement Not Met do
    POP' ← REPRODUCE(POP)
    simp ← GET-BEST(POP')
    if EVALUATE(simp) > EVALUATE(sbst) then
        sbst ← simp
    end if
    POP ← POP'
end while

```

---

### 3.4.3 Ant Colony Optimisation (ACO)

ACO, as a member of the swarm intelligence methods [102] and model-based search methods [103], is inspired by the foraging behavior of ants between the ant colony and food sources [104]. The history of past search is explicitly used to derive every constructive low-level solution. Additionally, ACO adopts a population framework and Monte Carlo style randomisation for a high-level selection based on the solutions achieved in the past [105].

Over time, a variety of ACO variants have been developed including: Ant System (AS) [104, 106, 107], Elitist AS [104, 107, 108], Ant-Q [109], Ant colony system [110], *Max – Min* AS [111, 112], Rank-based AS [113], Approximate Non-deterministic Tree-Search (ANTS) [114], Hyper-cube AS [115].

ACO can be applied to any combinatorial problem as each artificial ant constitutes a constructive procedure to define a complete solution by adding each component into the partial solution. A construction graph  $G = (V, E)$  is created where the edges set  $E$  connects each element in the vertex  $V$ . Each ant is capable of dealing with the constraints  $\Omega$  during its constructive procedure. Each edge  $e_{i,j} \in E$  between two vertices  $v_i$  and  $v_j$  is associated with two kinds of information: pheromone information  $\tau_{i,j}$  and heuristic information  $\eta_{i,j}$ . The  $\tau_{i,j}$  as a long-term memory stores the searching history of the artificial ants and is updated iteratively. The heuristic information  $\eta_{i,j}$  is a metaphor of the local sight of an artificial ant which is proportional to the cost of edge  $e_{i,j}$ . The heuristic information specifies the a priori information about the problem instance or run-time information provided from the source [84, p.36] and it will usually remain unchanged during the searching procedure.

The properties of the artificial ant are summarised as follows [84, p.36]:

- Each artificial ant exploits the construction graph  $G = (V, E)$  by a constructive procedure.
- Each artificial ant is capable of memorising the elements which have already been constructed so far. The functions of memory can be summarised as follows:
  - To deal with constraints;
  - To obtain the heuristic information;
  - To evaluate a complete solution after the construction is completed;
  - To retrace back the path to update the corresponding pheromone information.
- Before the termination conditions are satisfied, an artificial ant will select the next vertex in the neighbourhood of the current vertex through a probabilistic function which consists of:

- Local pheromone information and heuristic information;
  - Individual memory of the current constructive state of the ants;
  - Problem constraints.
- The pheromone associated with any edge passed by an artificial ant can be updated during the construction procedure or afterwards. The update procedure after construction is accomplished by retracing the same path backwards using the local memory of the artificial ant.
  - Each ant works concurrently and independently and communicates through mediating the information read or written by each ant.

There are three major functions to be defined in the ACO implementation process [84, p.37].

- **CONSTRUCTION** Within the construction graph of the problem  $G = (V, E)$ , a set of ants construct each respective solution by visiting each vertex in the current neighbourhood consecutively until a complete solution is constructed. Each step is determined by applying a probabilistic function mentioned in the property of an artificial ant. In some variants of ACO, the construction is in association with some local search techniques to promote the search performance. This function will return a set of feasible solutions, denoted by  $s$ .
- **DAEMON-ACTION** In some variants of ACO, this procedure controls the disposal of pheromone based on the information collected from local or global information. With the addition of such a procedure, the UPDATE procedure will be more purposefully attended in the hope of directing the search to a better solution more quickly. Pheromone set  $\Gamma$  is used to denote all the pheromone information of the construction graph. This optional function will take the responsibility of increasing the pheromone information of some of the edges.



- **UPDATE** The update procedure will increase or decrease the pheromone information associated with each edge. Generally, the update procedure will affect the popularity of each edge in the total construction of the solution. A solution with better quality will tend to increase the pheromone information to attract more ants to follow the track. Meanwhile, the reduction of pheromone will decrease the popularity of corresponding edges with the purpose of “forgetting” (or as referred to as evaporating [84]) any bad or local optimum solution.

---

**Algorithm 3.2** Pseudo-code for ACO

---

```
while Termination Conditions Not Met do  
     $s = \text{CONSTRUCTION}(\Gamma)$   
     $\Gamma = \text{DAEMON-ACTION}(\Gamma, s)$  {Optional}  
     $\Gamma = \text{UPDATE}(\Gamma, s)$   
end while
```

---

#### 3.4.4 Summary

In this section the general characteristics of metaheuristics are introduced. According to various dimensions of classification, metaheuristics can be categorised into different types. To summarise, to implement a metaheuristic several factors should be included, which are listed as follows:

- Initially, metaheuristics should produce a feasible solution randomly or heuristically (local search based) or constructively (constructive search);
- An evaluation function relevant to the objective function is essential;
- It is capable of obtaining the new feasible solution through a neighbourhood function or other mechanisms, e.g. pheromone;
- It has the criteria for selecting and accepting the solutions;
- It should be equipped with termination criteria.

## 3.5 Summary

In this chapter, some optimisation techniques are reviewed, including numerical optimisation algorithms, dynamic programming, and some metaheuristics. An introduction of a typical algorithm of numerical optimisation are covered in Section 3.2. Line search and trust region strategies are two of the important strategies to generate new solutions in numerical optimisation. Dynamic programming, due to its special problem solving characteristics, is illustrated ranging from the mathematic presentation to the comparison between dynamic programming and greedy algorithm. Finally, to tackle “NP-complete” problems, metaheuristics are introduced. The general characteristics of metaheuristics are covered and the elements essential to implementing a metaheuristic are presented.

## **Chapter 4**

# **Modelling train motion and traction power**

### **4.1 Introduction**

The application of computer simulation in railway engineering dates back to the days when computer techniques began to flourish in 1970s. The modelling of railway vehicle motion and traction power is no exception [116, 117]. Modelling of railway vehicle motion and traction power is done to calculate the speed, distance and time profile when a train is travelling on a journey with other constraints imposed on it, e.g. the signalling system and traction equipment characteristics. The signalling system is important in multiple train movement simulation because it controls the interaction between movements of the trains on a connected route [118, 119]. In [120] it is stated that train movement simulation can be applied to assess traction performance under the conditions of a new line [121], to verify a range of signalling systems [119, 122] and to evaluate proposed timetables [123].

In [116], different train movement models are reviewed. The time based simulation model evaluates train movement at every time interval and the state of the train is updated by assuming that its acceleration is constant in that interval [118]. It is also argued that, despite the fact that time-based modelling is conceptually close

to the practical train movement, the higher computational complexity limits its application to areas where the full details of every movement or energy consumption calculation are absolutely necessary [33]. What is not mentioned in [116], but has very similar modelling, is distance based modelling, which has been applied and reported in various papers [33, 121]. Both time and distance based modelling will be discussed in detail in Section 4.3. Different from time and distance based modelling techniques, without focusing on the details during the time of update, event based modelling can eliminate unrelated simulation and reduce the computational time. The movement of the train is denoted by a sequence of pre-defined events, such as arrival and departure at the stations [124]. A model of train movement is discussed in [49], used for the study of power strategy management of a DMU train.

## 4.2 Physics of vehicle motion

### 4.2.1 General introduction

Energy for the railway traction system is used for acceleration, overcoming electrical and mechanical power losses and for work in moving the mass of the train forward against the frictional forces [51].

The basic equation of motion is based on Newton's Second Law:

$$M^* \frac{d^2 s}{dt^2} = TE - R - Mg \sin(\alpha) \quad (4.1)$$

where:

- $M^*$  is the effective vehicle mass (including rotational inertia);
- $M$  is the vehicle mass;
- $g$  is the acceleration due to gravity;

- $R$  is the vehicle resistance to motion;
- $TE$  is the tractive effort;
- $s$  is the current vehicle distance.
- $\alpha$  is the angle between the route slope and vertical line.

### 4.2.2 Adhesion

Adhesion represents a limitation of the tractive effort produced by the motored axles and this should be considered before investigating the actual tractive force. The adhesion limit defined by Eqn. 4.2 sets the maximum available tractive effort  $TE_{max}$  on a flat track.

$$TE_{max} = \mu Mg \quad (4.2)$$

In Eqn. 4.2,  $TE_{max}$  is the maximum tractive effort;  $\mu$  is the coefficient of friction;  $M$  is the vehicle mass;  $g$  is the acceleration due to gravity. If  $\mu$  drops to too low a level, undesirable wheel slip may occur, causing energy wastage and preventing the vehicle from increasing its speed [55].

In [125], rail adhesion analysis using a full scale roller rig has been undertaken. The test results show the following:

1. For a dry and clean wheel surface, the adhesion coefficient  $\mu$  does not vary with vehicle speed;
2. For a track surface covered by oil, the adhesion coefficient  $\mu$  can drop down to a very low level and this level is approximately maintained with increase in vehicle speed;
3. For water covering the wheel/rail contact surface, the adhesion coefficient  $\mu$  reduces with increase in vehicle speed.

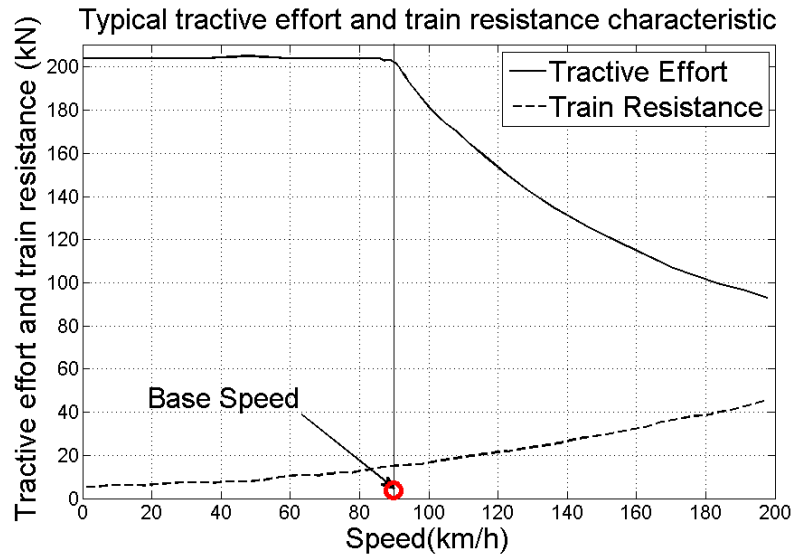


Figure 4.1: Typical railway vehicle tractive effort and resistance characteristic

It is also stated that the application of sand to the wheel/rail contact has been found to be the most efficient and cost effective method of increasing the adhesion coefficient  $\mu$  under contaminated conditions [126].

At higher speeds, power becomes the main limiting factor. In Figure 4.1, the vehicle keeps a constant tractive effort up to a particular point, referred to as the “base speed” and then has a constant power region which has the effect of decreasing the tractive effort reciprocally. Positive acceleration is maintained up to the balancing speed where the installed tractive capacity is equal to the combined resistance and gradient forces.

The frictional force between the driven steel wheels and the rail is the actual tractive force which moves the train forward. The value of the adhesion force on one driven axle can be calculated using Eqn. 4.2, where  $\mu$  is the coefficient of friction. In railroad applications, this value ranges from 0.05 to 0.5, depending on the conditions. The upper limit of 0.5 is sometimes assumed when encountering dry or sanded track [127].

In order to maximise the total tractive force, it is usual to increase the number of motored axles. The maximum acceleration on flat track is expressed by:

$$\frac{d^2s}{dt^2} = gp_m\mu \quad (4.3)$$

where:

- $g$  is the acceleration due to gravity;
- $p_m$  is the ratio of number of motored axles over total number of axles;
- $\mu$  is the coefficient of friction.

The coefficient of friction is assumed to be independent of the speed of the train, but in reality there is some decrease at high vehicle speeds.

### 4.2.3 Resistance

The motion of the train is opposed by a number of resistive forces [128, 129]. The overall resistance on level track can be formulated as follows:

$$\begin{aligned} R &= (A' + B'v)M + Cv^2 + \frac{D}{r} \\ &= A'M + B'Mv + Cv^2 + \frac{D}{r} \end{aligned} \quad (4.4)$$

Where  $A'$ ,  $B'$ ,  $C$  and  $D$  are empirical constants relating to rolling resistance, track resistance and aero-dynamics resistance,  $M$  is vehicle mass,  $v$  is the speed and  $r$  is the radius of track curvature. The constant terms in Eqn. 4.4 are commonly obtained using empirical methods [129]. Eqn. 4.5 is a simplified version that is used in the current work. Note that  $A = A'M$ ,  $B = B'M$ . The curvature of railway as represented by term  $\frac{D}{r}$  takes very limited effects when the train is running at speed less than 200 km/h. In this equation, the constants include the effect of mass and the increase in resistance due to track curvature has been assumed to be negligible.

$$R = A + Bv + Cv^2 \quad (4.5)$$

Where  $v$  is the vehicle speed ( $m/s$ ) and the coefficients  $A(kN)$ ,  $B(kNs/m)$  and  $C(kNs^2/m^2)$  are all constants, referred as the Davis coefficients [129].

### 4.2.4 Effective Mass

The rotational inertia of the rotating components on the train must be taken into account in order to properly calculate the acceleration of the train. This is usually done by adding a rotary allowance term to the mass of the train. This is expressed as a fraction of the tare mass of the train which is the mass of train without loads or passengers.

$$M^* = M_t(1 + \lambda_w) + M_l \quad (4.6)$$

Where:

- $M^*$  is the effective mass;
- $M_t$  is the tare mass;
- $\lambda_w$  is the rotary allowance;
- $M_l$  is the freight or passenger load.

Alternatively, for the sake of simplicity, the effective mass could be simply expressed by:

$$M^* = M \times (1 + \lambda_w) \quad (4.7)$$

The rotary allowance  $\lambda_w$  is a constant (usually less than 0.2) and depends on the ratio of number of motored axles over total number of axles, the gear ratio and type of vehicle construction [127].



### 4.2.5 General vehicle motion equation

To summarise, the general equation of vehicle motion, known as Lomonosoff's equation, can be written as follows:

$$M^* \frac{d^2 s}{dt^2} = TE - (A + B \frac{ds}{dt} + C \frac{d^2 s}{dt^2}) - Mg \sin(\alpha) \quad (4.8)$$

Here  $TE$  is the tractive effort and  $A$ ,  $B$ , and  $C$  are the constants in Eqn. 4.5 and  $-Mg \sin(\alpha)$  indicate the effect of the gradients.  $\alpha$  will be positive for uphill running, negative for downhill running and zero for flat track running.

In view of Eqn. 4.8, the instant acceleration  $a$  in relation to the instant speed  $v$  is given in Eqn. 4.9:

$$a = \frac{TE}{M^*} - \left( \frac{A}{M^*} + \frac{B}{M^*} v + \frac{C}{M^*} v^2 \right) - \frac{M}{M^*} g \sin(\alpha) \quad (4.9)$$

## 4.3 Modelling and simulation

The acceleration of the train, i.e. the right terms of Eqn. 4.8 over the effective mass of the vehicle, i.e.  $M^*$  shown in Eqn. 4.8, is determined by the following four issues [54].

- **Traction equipment specification** This includes the type of motors with their own control techniques. Different types of motor will result in different dynamic performance, for example, the initial acceleration at low speed, the maximum speed it could operate with, and its motoring and braking characteristics. The latter characteristic dominates the regeneration capability of the vehicle.
- **Mechanical design issues** These issues cover different aspects of the mechanical design of a train, such as weight, capacity, length, number of axles, proportion of motoring axles, train rolling resistance, gear ratios etc.

- **Topographic issues** These issues cover the gradient, curvature, position of station locations, and any other environmental issues such as weather conditions.
- **Operational control issues** The train should be operated under strict control to maintain satisfactory performance. Average speed restrictions and station dwelling times are among the control restrictions for safety considerations. For better passenger comfort, limits are also imposed on acceleration, jerk etc.

The first three issues remain unchanged for the same type of vehicle and routes. Operational control issues are usually the control input for instantaneous train movement. Consequently, the motion of the single railway vehicle is generally deterministic regardless of random disturbances from the outside environment [54].

#### 4.3.1 Vehicle state switch

Let vehicle state  $\phi$  present the combination of vehicle current distance  $S_{cur}$ , current journey time consumed  $T_{cur}$  and current vehicle speed  $V_{cur}$ . The distance switch from the initial state to the final one results in the train movement from one station to the next. The tractive power from the locomotive engines or traction motors is the internal cause of the state switch. Different power imposed on the vehicle current state will result in a different vehicle state switch, i.e. train running trajectory. This section will illustrate the physical idea behind the vehicle state switch and consequently the idea of modelling the vehicle motion and its simulation.

For a known vehicle state, to obtain the next state after the state switch, the assumption should be made that: “the acceleration imposed on the vehicle should be known and usually regarded as constant within a certain time interval or within a certain distance interval.” The time interval or distance interval are kept short to keep certain level of calculation precision. Since the acceleration imposed on the vehicle is comprised of the non-linear resistance, gradients and tractive effort from

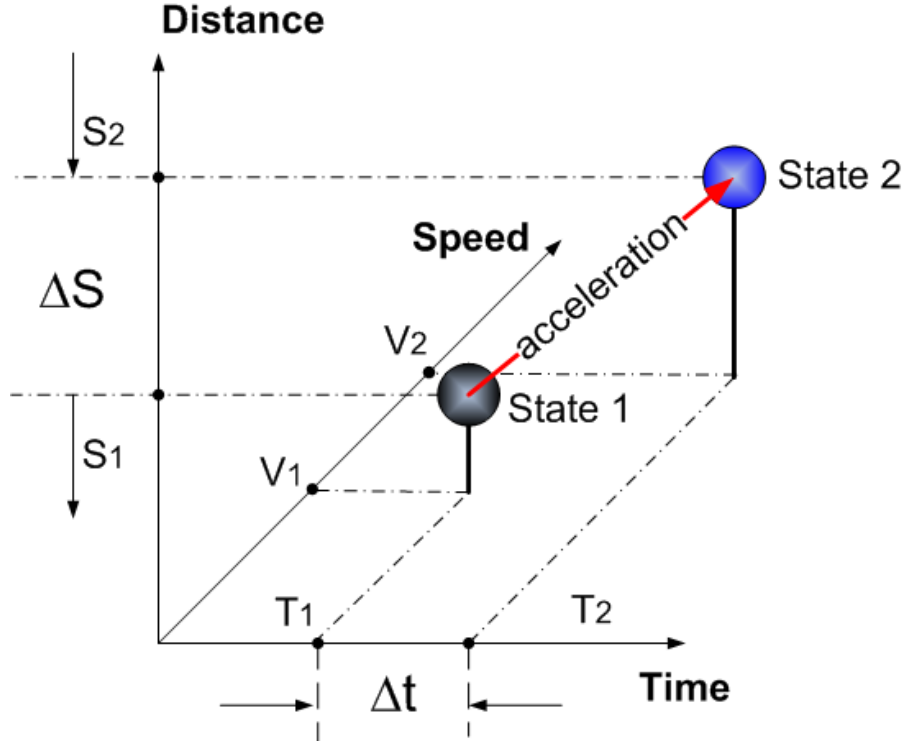


Figure 4.2: Vehicle state switch between two states in 3 dimension space

tractive equipment, the numerical solution for the vehicle motion can be obtained based on this assumption.

Figure 4.2 shows a simple case of a state switch between two vehicle states. Due to the constant positive acceleration in the period  $\Delta T$ , the speed will be changed from  $v_1$  to  $v_2$  and distance will be changed from  $s_1$  to  $s_2$ . If the instant acceleration is  $a$  and it is regarded as constant within the next period  $\Delta T$  (time based modelling) or next distance interval  $\Delta S$  (distance based modelling), the state 2 originating from a known state 1 is obtained.

For time based modelling, for which acceleration is regarded as constant within a time interval for a simulation step, the process of obtaining state 2 can be illustrated by Eqn. 4.10 and Eqn. 4.11. In this case,  $\Delta T$  is a known constant.

$$v_2 = v_1 + a \cdot \Delta T \quad (4.10)$$

$$s_2 = s_1 + \frac{v_2^2 - v_1^2}{2 \cdot a} \quad (4.11)$$

Time based modelling can simulate the train motion system as a dynamic time based system [49] and, since the vehicle states can be compared using the same discretised time basis, it is easier to apply DP to obtain the optimum vehicle running trajectory.

For distance based modelling, where the acceleration is regarded as constant within a distance interval for a simulation step, the process of obtaining state 2 can be illustrated by the following equations. In this case,  $\Delta S$  is a known constant.

$$v_2 = \sqrt{v_1^2 + 2 \cdot a \cdot \Delta S} \quad (4.12)$$

$$t_2 = t_1 + \frac{v_2 - v_1}{a} \quad (4.13)$$

Distance based modelling has the advantages of a stable distance basis, thus it is easier to store the data and enable a faster search metaheuristic for a good vehicle running trajectory.

Because the initial vehicle state is known, i.e.  $s = 0$ ,  $t = 0$ , and  $v = 0$ , if at each simulation step the instant acceleration  $a$  is also known, all the states along the journey are determined. Figure 4.3 shows the state switch along a simple journey using different operation modes.

### 4.3.2 Operational control input

The operational control input can be generally categorised into the four types: motoring, cruising, coasting and Braking.

#### Motoring mode and cruising mode

This control will increase the vehicle speed and switch the vehicle from a low speed state to a higher speed state and keep at a constant level under cruising mode. The acceleration under such control will be positive or zero, i.e.  $a \geq 0$ .

For a vehicle with a known state  $\phi$ , with current speed and current distance, the

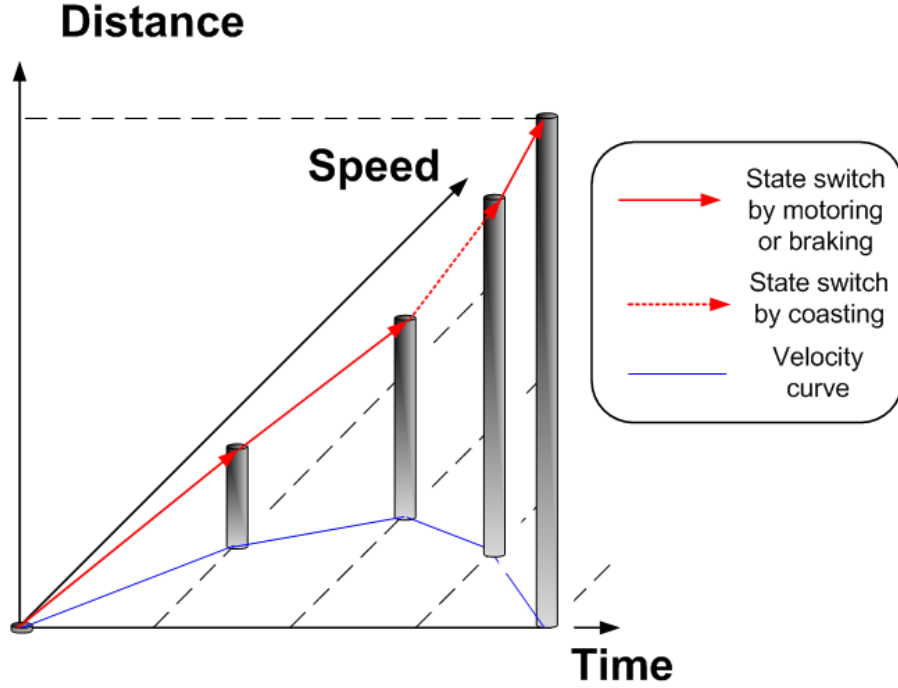


Figure 4.3: Vehicle states switch in 3 dimension space

current gradient  $Mg \sin(\alpha)$ , current resistance  $R$  in Eqn. 4.1 is known and regarded as stable for the current simulation step. The instantaneous maximum operational acceleration  $a_{max}$  is determined by the maximum achievable acceleration  $a_{max}^{ach}$  determined by the maximum tractive effort from the tractive equipment shown in Figure 4.1 and also by the current speed limit. As long as acceleration is kept positive, the motoring mode remains active.

To simplify the illustration, a train is assumed to accelerate to speed limit using maximum available acceleration. To determine the maximum acceleration from the current maximum tractive effort and speed limit, a speed threshold  $v_{th}$  is set for current speed limit  $v_{lim}$ , (Note: in this thesis  $v_{th} = v_{lim} - C$ , where “ $C$ ” is a small constant value in m/s) and the following two rules are proposed to be followed in the simulation to achieve a gradual approaching to the speed limit.

- Above the threshold, the maximum achievable acceleration  $a_{max}^{ach}$  will be truncated using Eqn. 4.15 to obtain the  $a_{max}$ ;
- Below the threshold speed,  $a_{max}^{ach}$  equals  $a_{max}$ .

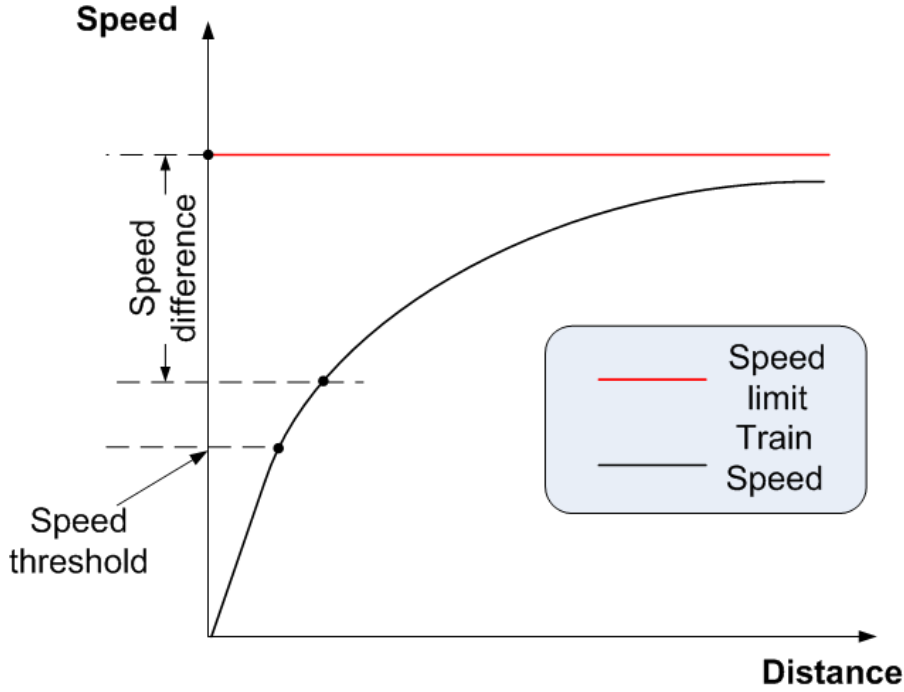


Figure 4.4: Maximum achievable acceleration truncation using partial open loop control

The truncation of  $a_{max}^{ach}$  is done using Eqn. 4.15.  $V_{diff}$  is to denote the current speed positive difference between the  $V_{cur}$  and  $V_{lim}$  as shown in Eqn. 4.14.

$$v_{diff} = v_{lim} - v \quad (4.14)$$

The maximum acceleration can be calculated using Eqn. 4.15.

$$a_{max} = a_{max}^* \cdot \frac{V_{diff}}{V_{lim} - V_{th}} \quad (4.15)$$

When the  $v_{diff} = 0$ ,  $a_{max} = 0$  and the train enters the cruising mode operation.

No matter which type of modelling is used, at each simulation step, the control input  $a_c$  for the vehicle is obtained from  $a_{max}$  and  $\lambda_c$ . The control input  $a_c$  is calculated using Eqn. 4.16 given that  $\lambda_c \in [0, 1]$ .

$$a_c = \lambda_c \cdot a_{max} \quad (4.16)$$

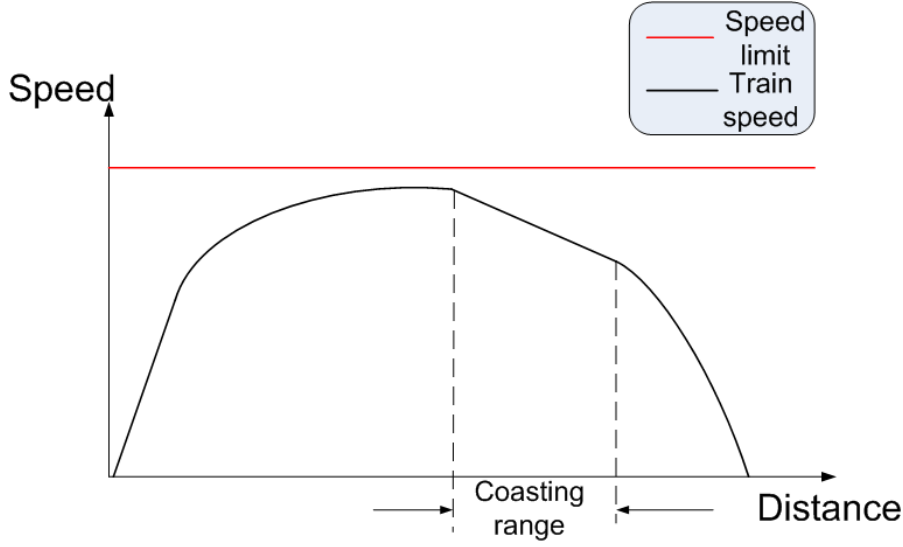


Figure 4.5: Coasting mode operational control

### Coasting mode

Coasting mode operational control will shut the motors or any tractive equipment down. Acceleration will be affected by the total resistance of the train and the effect of gravity. The acceleration  $a$  of the vehicle is usually negative, however, it can become positive if the vehicle is running on a steep downhill gradient.

The significance of coasting mode running is that the motors are using no energy during this mode. For non-regenerative braking operation, such operational control would save the most significant amount of energy. However, the selection of coast point from which the train is to begin to coast will have a significant effect on the both total journey service quality in terms of journey time and energy consumption [24, 25]. Figure 4.5 presents the speed curve under a simple coasting mode operational control strategy.

For both distance and time based modelling, the mechanism of simulation under coasting mode operational control is the same. In simulation, the instantaneous coasting acceleration  $a$  for each state is calculated by setting the  $F = 0$  in Eqn. 4.9.

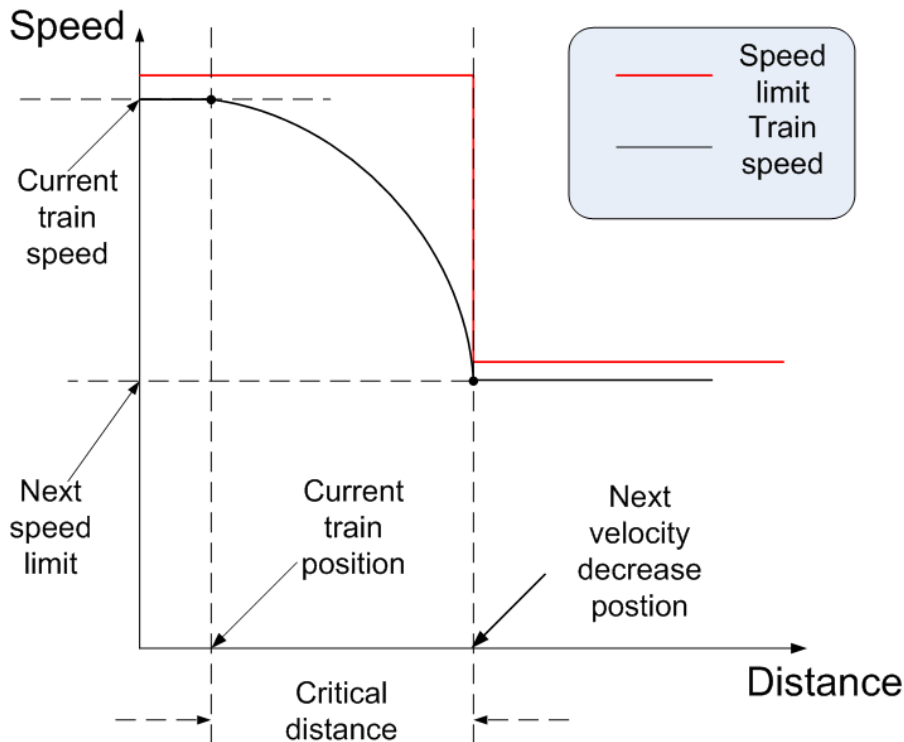


Figure 4.6: Braking mode operational control due to reduced speed limit

### Braking mode

Braking mode slows the train down to meet speed limits and also to stop at stations. The critical requirements of braking operational control are to make sure that the speed of the train is always under the speed limit and that the train arrives at the station with zero speed.

For time based modelling, the braking mode of the train is assumed to be a service braking procedure with a typical constant deceleration rate. According to the function, braking mode operational control can be categorised into two types: type I, braking due to the decrease of the speed limit, and type II, braking due to the station stop. Figure 4.6 demonstrates a speed reduction procedure due to the decrease of the speed limit.

Figure 4.7 demonstrates a speed reduction procedure due to a station stop.

The braking mode operational control is not taken until the “Critical Braking Distance (CBD)” is reached. CBD is determined by Eqn. 4.17 for type I braking,



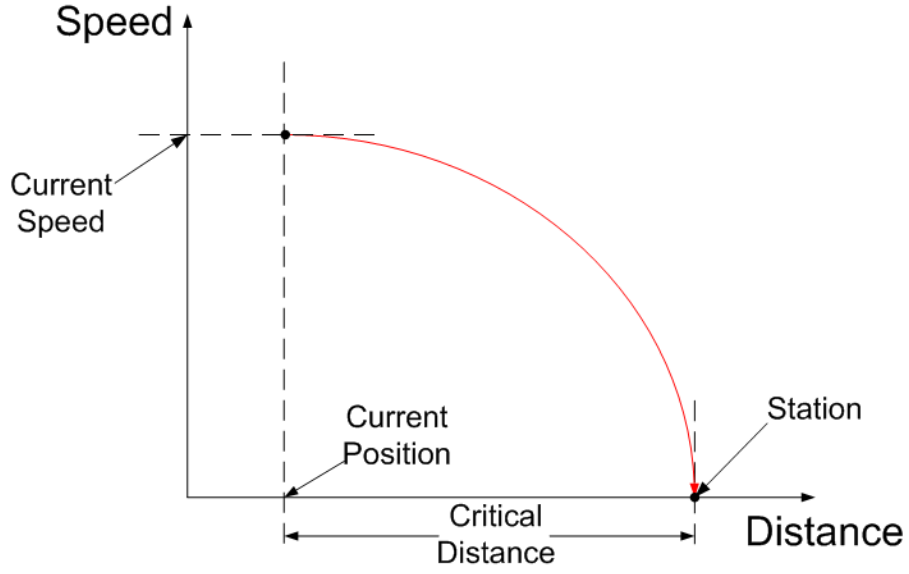


Figure 4.7: Braking mode operational control due to station stop

and by Eqn. 4.18 for type II braking.

$$s_c = -0.5 \cdot (v_{nl}^2 - v^2) / a_{svc} \quad (4.17)$$

$$s_c = -0.5 \cdot v^2 / a_{svc} \quad (4.18)$$

In Eqn. 4.17 and Eqn. 4.18,  $s_c$  is the CBD;  $v_{nl}$  is the next speed limit;  $s_c$  is the current speed for the current vehicle state;  $a_{svc}$  is the presumed constant service acceleration.

For distance based modelling, the braking mode operational control is simulated by a virtual backwards calculation [33, 121]. It is assumed that an identical train is running backwards from the next station or next speed limit position with the same amount of braking effort plus dragging resistance calculated using Eqn 4.5. The deceleration rate of forwarding train is equal acceleration rate of backwards train. As the braking effort is now in the direction of the train, the speed change at each state should be the same as the one going forward. The backwards running train will keep running until its speed exceeds the forward speed trajectory, shown as braking position in Figure 4.8.

It is noted that for the braking mode, there is a considerable difference be-

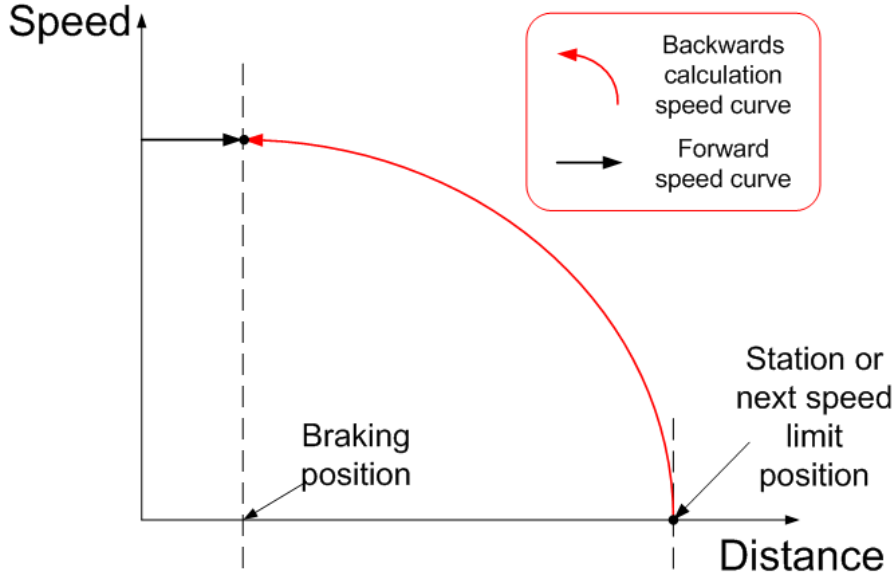


Figure 4.8: Braking mode operational control at distance based modelling

tween distance based and time based modelling. For time based modelling, it is not straightforward to obtain the braking speed for each simulation step. The braking procedure is approximated by introducing a constant service braking deceleration rate. By contrast, it is easier for the distance based simulation model to virtually calculate the braking speed backwards.

### 4.3.3 Energy consumption modelling

For each state switch, the instant acceleration of the vehicle  $a$ , is known based on the current information including the driver control input and the other external conditions, shown in Eqn. 4.9.

The energy consumed by the traction system of a train can be calculated using Eqn. 4.19.

$$E = \frac{1}{2}M(v_2^2 - v_1^2) + \int_{v_1}^{v_2} (Av^2 + Bv + C)ds - \Delta h * M * g \quad (4.19)$$

In Eqn. 4.19,  $M$  is the total vehicle mass,  $v_1$  is the vehicle speed at the current distance step and  $v_2$  is the vehicle speed at the next distance step,  $\Delta h$  is the altitude

difference between the altitude of the current distance step  $A_1$  and the altitude of the next distance step  $A_2$ , i.e.  $A_1 - A_2$ ,  $g$  is the acceleration due to gravity and  $A$ ,  $B$  and  $C$  are the Davis coefficients shown in Eqn. 4.5.

If the acceleration is constant during the speed change from  $v_1$  to  $v_2$  as assumed, the instantaneous speed at each distance from  $s_1$  to  $s_2$  in Figure 4.2 can be calculated by Eqn. 4.20.

$$v = \sqrt{v_1^2 + 2a\delta s} \quad (4.20)$$

In Eqn. 4.20,  $\delta s$  is the distance between distance  $s$  to  $s_1$ . If both  $v_1$ ,  $v_2$  are known, the constant acceleration  $a$  can be obtained from Eqn. 4.21.

$$a = \frac{v_2^2 - v_1^2}{2\Delta s} \quad (4.21)$$

In view of Eqn. 4.21, Eqn. 4.20 and Eqn. 4.19, the final traction energy cost for the train moving from  $s_1$  to  $s_2$  with the speed change from  $v_1$  to  $v_2$  can be presented as follows.

$$E = \frac{1}{2}M(v_2^2 - v_1^2) + A\Delta s/2(v_2^2 - v_1^2) + \frac{2B\Delta s(v_2^2 + v_2 \cdot v_1 + v_1^2)}{3(v_2 + v_1)} + Av_1^2\Delta s + C\Delta s - \Delta hMg \quad (4.22)$$

The time consumption can be calculated from Eqn. 4.23.

$$\begin{aligned} T &= \frac{2(v_2 - v_1)(s_2 - s_1)}{v_2^2 - v_1^2} \\ &= \frac{2(s_2 - s_1)}{v_2 + v_1} \end{aligned} \quad (4.23)$$

#### 4.3.4 Single train motion simulator

A Single Train Motion Simulator (STMS) is developed using Lomonosoff's Eqn. 4.8. An overview schematic is shown in Figure 4.9.

In Figure 4.9, four parts of the simulator are shown. From left to right, they are:

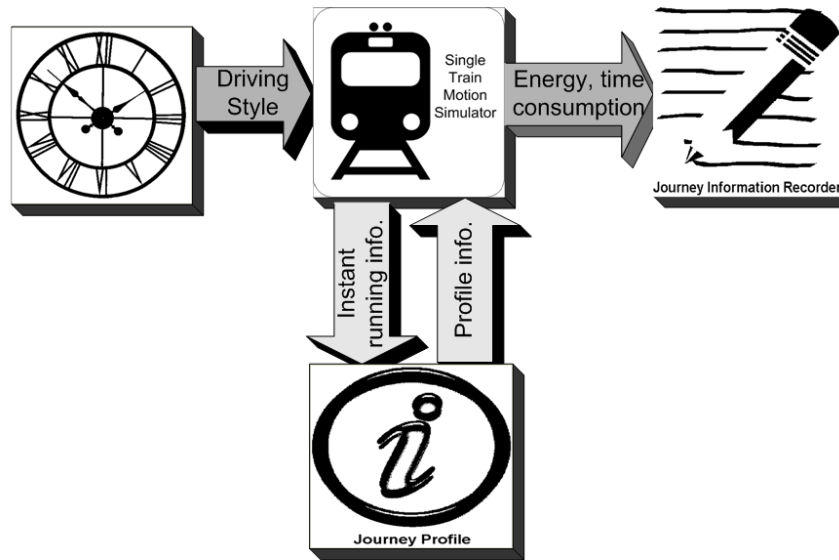


Figure 4.9: System diagram of single train simulator

1. **Journey Profile.** The journey profile, including the signalling system, geographic information, speed limit and station position, are all determined for a known journey. In the simulation, journey profiles are stored in a look-up table which is loaded before the simulation.
2. **Operational Control Input.** The STMS obtains the acceleration of the train from the data stored in “Journey Profile” and the control input from the driver as it moves through the journey at each simulation step.
3. **Journey Information Recorder.** To demonstrate the simulation results, the output results are stored in the Journey Information Recorder. At each time step, the states of the vehicle will be stored, including the distance of the vehicle, the speed of the train, and the energy and time consumed thus far.
4. **Single Train Motion Simulator.** By calculation, this block obtains the state switch based on the current state, and so on, until the final state is achieved.

## 4.4 Summary

This chapter contains a literature review of the modelling of a traction power system. Modelling generally follows laws of physics and logic and predicts the behaviour and performance of railway infrastructure [130]. As a result, establishment and development of correct a mathematical model of traction power systems is vital and worth detailed investigation. Simulation based on correct modelling provides a cost effective evaluation of a large range of energy efficiency studies on traction systems such as power management strategies and driving strategies. The physics of vehicle motion are introduced in Section 4.2. Each of the elements in Eqn. 4.8 is illustrated in detail. Key factors in modelling a traction power system are covered in Section 4.3. Specifically, the energy consumption calculation is discussed in Section 4.3.3.

# Chapter 5

## Single train trajectory optimisation

### 5.1 Introduction

Single train speed trajectory, i.e. the curve representing the speed profile, achieved by improved train traction control strategies can potentially reduce the energy consumption of railway vehicles. Driver guidance systems [131] or Automatic Train Operation (ATO) systems [132] are able to take advantage of pre-calculated train speed trajectories. Train speed trajectory optimisation has already been studied using various methods, such as GA and optimal control theory. Generally, train speed trajectory optimisation can be categorised into two types: coasting control optimisation and general control optimisation.

Coasting control optimisation optimises the train trajectory through coasting operations to ensure minimum energy consumption under the journey time constraints. The optimisation procedure regards the coasting operation as an optimisation input for a good tradeoff between energy consumption and journey time. The GA has also been applied to search for the coasting points but the number of coasting points is predetermined which limit its application for different journey profiles [25]. The results show a promising performance in coasting control for the tradeoff between the journey time and energy consumption. In [133], GA is also applied to search for the coasting point. The number of coasting points has been

dynamically allocated into the chromosomes which enhance its practical application. In [24], in addition to the GA searching methods some of the classic searching methods, i.e. golden search methods, are studied in a simple single coasting point case. Rather than search for the coasting point, [26] targets the acceleration rate, the braking rate and the re-motoring speed. In this paper the coasting operation is regarded as the natural operation after the train has reached the maximum speed limit. In addition, Artificial Neural Network (ANN) have been applied to search for the optimum coasting speed to minimise the total “social cost” in a mass rapid transit system [27]. In [27], the coasting control means the speed from which the train begins to coast. It is found that coasting speed should increase when more emphasis is placed on the journey time.

General control optimisation seeks the optimal train trajectory by applying a variety of control means, i.e. acceleration, deceleration, coasting, cruising etc. This type of search obtains more feasibility and controllability for practical applications as the control inputs are echoed by real train operations. General control includes the train running acceleration rate, deceleration rate and cruising speed level in addition to the coasting control to search for the optimum trajectory. Optimal control theory is among the most widely applied techniques used to obtain the optimal control sequence of the train. The optimisation objective is to operate a train with minimum energy consumption subject to time and physical limitations. The analytical solution of the problem is obtained through some linear approximation or some empirical extensions. Pontryagin Maximum Principle (PMP) is a commonly used method to achieve the solution [19, 29, 30]. DP has been applied to search for the optimum trajectory with the minimum energy cost [31, 32]. In [33], a multi-population genetic algorithm together with heuristic annealing selection has been modelled to search for an energy-saving control for an urban railway train. It is alleged that a multiple population search improves the convergence rate and evolution stability. In [123] a heuristic algorithm is proposed where the target speed is pre-obtained. A series of feasible control inputs is then produced. The simulation

results show that the train is able to recover from delay by using this algorithm.

In this chapter, train trajectory optimisation will be discussed firstly from an optimal control view-point. It is concluded that the optimisation of train trajectory can be also realised through the search of the train speed at various positions. A distance based model is created onto which several general control optimisation algorithms are applied to search for the optimum train running trajectories. ACO as the representative of constructive heuristic algorithms, and a GA representing the local search based metaheuristic, will both be applied to search for a more desirable train trajectory. For comparison study purposes, DP as a global optimisation technique is also applied.

## 5.2 An optimal control view-point

### 5.2.1 System plant

The train trajectory is manipulated by a train control input from the driver or any ATO system. The objective function for the optimal control can be presented by Eqn. 5.1:

$$J = \int_0^T u_f(t) f[v(t)] s(t) dt \quad (5.1)$$

where:  $u_f$  is the control signal for forward tractive effort,  $f[v(t)]$  is the maximum tractive effort per unit mass at current speed  $v(t)$ ,  $s(t)$  is the current vehicle distance. It is assumed that for backwards traction, i.e. braking operations, no power is consumed, thus no cost will be induced by a backwards control signal.

The train motion control is a time-variant two-point boundary value problem if the arrival time is fixed. We set the time as the dependant variable. The state equations are given by Eqn. 5.2.

$$\begin{cases} \dot{s} = v \\ \dot{v} = u_f \cdot f(v) - u_b \cdot b(v) - r(v) - g(s) \end{cases} \quad (5.2)$$


---



where,  $u_b$  is the control signal for backwards braking effort,  $b(v)$  is the maximum braking effort per unit mass at current speed  $v$ ,  $r(v)$  is the resistive force per unit mass defined by Eqn. 4.5 and  $g(s)$  is the longitudinal force per mass due to gradient.

The boundary conditions are listed as follows.

$$v(0) = 0 \qquad \qquad \qquad v(T) = 0 \qquad \qquad (5.3)$$

$$s(0) = 0 \qquad \qquad \qquad s(T) = S_t \qquad \qquad (5.4)$$

Some constraints are imposed:

$$v \leq v_{lim}(s) \qquad \qquad \qquad (5.5)$$

$$u_f \in [0, 1] \qquad \qquad \qquad (5.6)$$

$$u_b \in [0, 1] \qquad \qquad \qquad (5.7)$$

where  $v_{lim}(s)$  is the speed limit at various positions,  $u_f = 1$  implies the maximum possible tractive effort is imposed,  $u_f = 0$  implies no tractive effort is imposed,  $u_b = 1$  implies the maximum possible braking effort is imposed and  $u_b = 0$  implies no braking effort is imposed.

### 5.2.2 The pontryagin minimum principle

Using the PMP, the Hamiltonian “ $H$ ” [134, 135] can be presented as follows:

$$H = p_0\{\lambda_s[v - v_{lim}(s)] + u_f \cdot f(v)v\} + p_1 \cdot v + p_2 \cdot [u_f \cdot f(v) - u_b \cdot b(v) - r(v) - g(s)] \quad (5.8)$$

where  $\lambda_s$  is the complementary slackness coefficient and  $p_0$  is a positive constant and  $p_1$  and  $p_2$  are the costates depending on time.

$$\lambda_s = \begin{cases} = 0 & \text{if } v \leq v_{lim}(s) \\ > 0 & \text{if } v > v_{lim}(s) \end{cases} \qquad \qquad (5.9)$$

Rearranging Eqn. 5.8, gives:

$$H = p_1 \cdot v - p_2 \cdot r(v) - p_2 \cdot g(s) + p_0 \lambda_s [v - v_{lim}(s)] + p_2 \cdot u_f \cdot f(v) - p_2 \cdot u_2 \cdot b(v) + p_0 \cdot u_f \cdot f(v) \cdot v \quad (5.10)$$

If  $x_1$  is used to represent  $s$ ,  $x_2$  to represent  $v$ ,  $u_1$  to represent  $u_f$  and  $u_2$  to present  $u_b$ , Eqn. 5.10 can be written as follows:

$$H(t, x, p, u) = I(t, x, p) + F(t, x, p) \cdot u \quad (5.11)$$

where,

$$I(x, p, t) = p_1 \cdot x_2 - p_2 \cdot r(x_2) - p_2 \cdot g(x_1) + \lambda_s \cdot p_0 [x_2 - v_{lim}(x_1)] \quad (5.12)$$

$$F(x, p, t) = [p_2 \cdot f(x_2) + p_0 \cdot f(x_2) \cdot x_2, -p_2 \cdot b(x_2)] \quad (5.13)$$

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (5.14)$$

In addition,  $F_1$  and  $F_2$  are used to denote the two elements of vector  $F$ .

$$F_1(x, p, t) = p_2 \cdot f(x_2) + p_0 \cdot f(x_2) \cdot x_2 \quad (5.15)$$

$$F_2(x, p, t) = -p_2 \cdot b(x_2) \quad (5.16)$$

Costate equations can be shown as follows:

$$\dot{p}_i = -\frac{\partial I}{\partial x_i} - u(t) \frac{\partial F}{\partial x_i} \quad (5.17)$$

Suppose that there is an optimal control vector  $u^*$  with its resultant the optimum trajectory  $x^*$  vector within the determined time range, according to PMP the following conditions must hold [134, 135]:

- All state functions described in Eqn. 5.2, Eqn. 5.3, Eqn. 5.4 and the costate

Eqn. 5.17 should hold. The equations are presented as follows:

$$\dot{x}_1^* = x_2 \quad (5.18)$$

$$\dot{x}_2^* = u_1^* f(x_2^*) - u_2^* b(x_2^*) - r(x_2^*) - g(x_1^*) \quad (5.19)$$

$$\dot{p}_1^* = -p_2 g'(x_1) \quad (5.20)$$

$$\begin{aligned} \dot{p}_2^* = & -u_1 f'(x_2) x_2 - u_1 f(x_2) - p_1 - p_2 u_1 f'(x_2) \\ & - p_2 u_2 b'(x_2) - p_2 r'(x_2) \end{aligned} \quad (5.21)$$

and the boundary conditions are:

$$x_1^*(0) = 0 \quad x_1^*(T) = S_t \quad (5.22)$$

$$x_2^*(0) = 0 \quad x_2^*(T) = 0 \quad (5.23)$$

- For  $t \in [0, T]$  and all the  $u(t)$  satisfying the condition  $|u(t)| \leq 1$ , the Hamiltonian takes its supremum at the optimal control  $u^*$  for all the admissible  $u$  set.

$$H(x^*, p^*, U^*, t) \leq H(x^*, p^*, U, t) \quad (5.24)$$

- The Hamiltonian  $H$  is a constant along the optimum trajectory in this case where the terminal time is fixed.

$$H(x^*, p^*, U^*, t) = c \quad \forall t \in [0, T] \quad (5.25)$$

where,  $c$  is a constant.

### 5.2.3 Singular control solution

It is noted that the relationship between the input control  $u$  and Hamiltonian  $H$  is linear and this is a problem where PMP fails to achieve a complete solution [135, 136]. The summaries for the control vector  $u$  are restated in this case [33].

1. If  $F_1 > 0$  and  $F_2 < 0$ , this gives  $u_1 = 0$  and  $u_2 = 1$ , corresponding to the full braking operational mode;

$$\begin{cases} F_1 = p_2 \cdot f(x_2) + p_0 \cdot f(x_2) \cdot x_2 > 0 \\ F_2 = -p_2 \cdot b(x_2) < 0 \end{cases} \quad (5.26)$$

$$p_2 > 0 \quad (5.27)$$

2. If  $F_1 > 0$  and  $F_2 > 0$ , this gives  $u_1 = 0$  and  $u_2 = 0$ , corresponding to the coasting operational mode;

$$\begin{cases} F_1 = p_2 \cdot f(x_2) + p_0 \cdot f(x_2) \cdot x_2 > 0 \\ F_2 = -p_2 \cdot b(x_2) > 0 \end{cases} \quad (5.28)$$

As a result:

$$\begin{cases} x_2 > -\frac{p_0}{p_2} \\ p_2 < 0 \end{cases} \quad (5.29)$$

As  $p_0$  is a positive constant, Eqn. 5.29 indicates that the coasting operation should not occur unless the vehicle speed is over a certain level.

3. If  $F_1 < 0$  and  $F_2 < 0$ , this gives  $u_1 = 1$  and  $u_2 = 1$ . This corresponds to a non-real operational mode where the speed is negative, shown in Eqn. 5.31.

$$\begin{cases} F_1 = p_2 \cdot f(x_2) + p_0 \cdot f(x_2) \cdot x_2 < 0 \\ F_2 = -p_2 \cdot b(x_2) < 0 \end{cases} \quad (5.30)$$

$$\begin{cases} x_2 < -\frac{p_0}{p_2} \\ p_2 > 0 \end{cases} \quad (5.31)$$

4. If  $F_1 < 0$  and  $F_2 > 0$ , this gives  $u_1 = 1$  and  $u_2 = 0$ , corresponding to full
-

power operational mode.

$$\begin{cases} F_1 = p_2 \cdot f(x_2) + p_0 \cdot f(x_2) \cdot x_2 < 0 \\ F_2 = -p_2 \cdot b(x_2) > 0 \end{cases} \quad (5.32)$$

Therefore:

$$\begin{cases} x_2 < -\frac{p_0}{p_2} \\ p_2 < 0 \end{cases} \quad (5.33)$$

It is noted that full power is not taken, for speed  $x_2$  becomes above a certain speed level, for example, when the train is reaching the speed limit less power should be adopted to prevent any overshooting the speed limit.

These four types of controls are referred to as “bang-bang control”, where the control signal takes only the boundary value within the allowable range [137]. When either  $F_1$  or  $F_2$  becomes zero, the singular control should be adopted to maintain  $I \equiv 0$  and  $F \equiv 0$ . Two definitions are given as below.

**Definition 5.1** (Optimal Singular Control). *The problem defined by Eqns. 5.2 to 5.7 is singular, if the optimal control  $u^*$ , the resultant train trajectory  $x^*$  and the costate  $p^*$  have the following property.*

*There is at least one half open interval  $(t_1, t_2]$  such that:*

$$F(x, p, t) = 0 \quad (5.34)$$

*The control function  $u^*_{(t_1, t_2]}$  is then called the singular control and the trajectory  $x^*_{(t_1, t_2]}$  an optimal singular trajectory.*

**Definition 5.2** (Extremal Singular Control). *If there is an extremal control  $\hat{u}$  which satisfies all the necessary conditions defined by PMP, such that the corresponding trajectory  $\hat{x}$  and the costate  $\hat{p}$  have the property that there is at least one open half*

time range within which the following equation holds.

$$F(\hat{x}, \hat{p}, t) = 0 \quad (5.35)$$

In practice, both the powering operation and braking operation could not exist simultaneously, one of the control signals will be set to zero to eliminate the effect of non-zero  $F$  elements. For example, if  $F_1 = 0$  and  $F_2 \neq 0$ ,  $u_2$  will be set to zero to eliminate the effect of non-zero  $F$ . In that case, the situations where  $F_1 = 0$ ,  $F_2 < 0$  or  $F_1 < 0$ ,  $F_2 > 0$  will not be considered.

1. If  $F_1 = 0$  and  $F_2 > 0$ , this gives  $u_1 \in [0, 1]$  and  $u_2 = 0$ , corresponding to the partial power operational mode.
2. If  $F_1 > 0$  and  $F_2 = 0$ , this gives  $u_1 = 1$  and  $u_2 \in [0, 1]$ , corresponding to the partial braking operational mode.

According to the two definitions, during the partial power operational mode for the time range  $(t_1, t_2]$ ,  $u_1 \in [0, 1]$  should keep  $F_1 \equiv 0$  to satisfy the PMP necessary conditions defined in Eqns. 5.17, 5.24 and 5.25. Meanwhile  $u_2 \in [0, 1]$  for the partial braking operational mode for  $(t_1, t_2]$  should keep  $F_2 \equiv 0$ .

It is not possible to obtain the analytical control input  $\hat{u}$  due to the non-linear characteristic of the real train running. Some reasonable approximation should be taken in order to search for the optimal control signal and resultant trajectories. For example, in [19], the energy cost of the journey is assumed to be linear to the journey time. However, such approximation does not always provide reasonable results due to the practical constraints. It is also not straightforward to apply the numerical method to obtain the solution using PMP due the singular characteristic. As a matter of fact, the “extremal singular control” does not necessarily imply an “optimal singular control” [135, 136]. This generally means the results obtained by PMP are not always optimal. In addition, the numerical method always has the disadvantage that the initial selection of the solution will severely affect the

searching performance to approach the final optimum answer as shown in Section 3.2.

However, the necessary conditions for optimal control do clearly give a group of categories of operations from which the optimal control could be selected in a certain sequence [29, 30].

To avoid the complexity of searching for the optimal control, without losing generality, this thesis assumes that the train running within a short enough distance with the same initial speed and same terminal speed will consume approximately identical energy and identical time. The assumption is made based on similar reason for numerical calculation of vehicle states in Section 4.3.1. The candidate operations in the optimal control will result in the different speed switch between the two positions. In this case, the optimisation procedure will be changed into search for the suitable speed at different positions. Meanwhile, to reflect the optimal control theory briefly discussed in this section, certain speed changes corresponding to certain potential optimal controls are included in the searching space, i.e. full powering operational mode, full braking operational mode, coasting operational mode, partial powering and partial braking operation modes.

## **5.3 Modelling context**

### **5.3.1 Introduction**

In Chapter 4, the modelling of both train motion and traction power is discussed in detail. This chapter adopts distance based modelling to simulate train motion and traction power due to the advantage of such modelling in graphical search. The motion of the train is calculated step by step in terms of distance. The target of train trajectory optimisation is to search for the speed for each position along the journey and try to find the minimum energy consumption meeting the punctuality

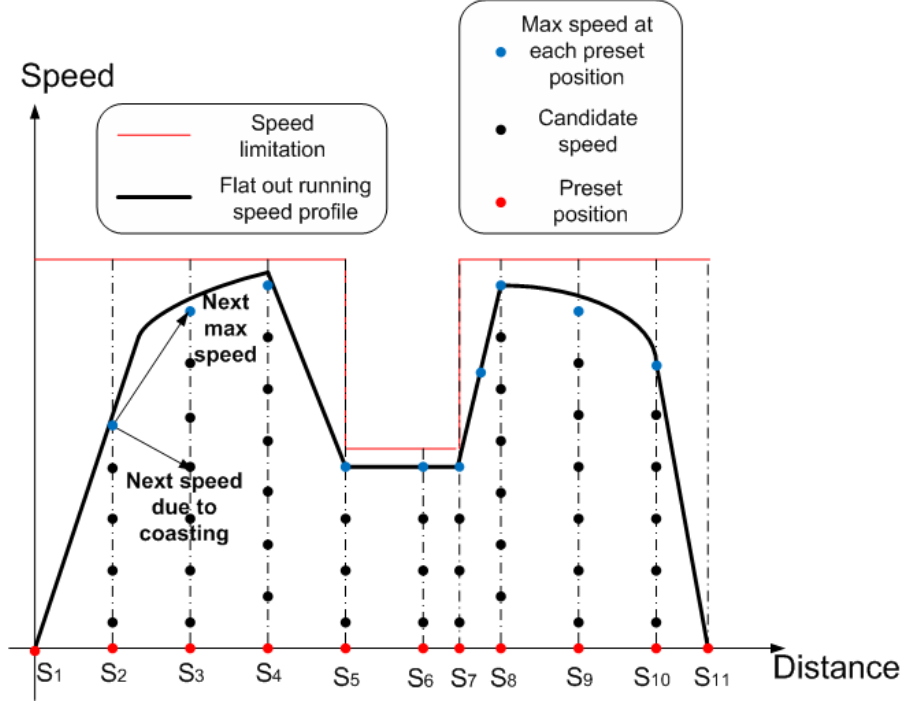


Figure 5.1: Speed selection procedure in the distance based trajectory searching

requirement. The evaluation function is defined as follows:

$$J_{tra} = \begin{cases} E & \forall T \in [T_{lb}, T_{ub}] \\ E + P & \forall T \notin [T_{lb}, T_{ub}] \end{cases} \quad (5.36)$$

where  $E$  is the energy consumption for the proposed journey trajectory,  $T$  is the time cost for the proposed journey trajectory, and  $P$  is the penalty function which outputs a positive number with an increasing value if the journey time either increases or decreases from the accepted journey time range defined by  $[T_{lb}, T_{ub}]$ .  $T_{lb}$  is the minimum allowable journey time and  $T_{ub}$  is the maximum allowable journey time. If  $T_{prop}$  denotes the proposed journey time, the following relationships hold.

$$T_{lb} \leq T_{prop} \quad (5.37)$$

$$T_{ub} \geq T_{prop} \quad (5.38)$$



In this thesis the  $T_{lb}$  is defined as 1% less than  $T_{prop}$  and  $T_{ub}$  is defined as 1% more than  $T_{prop}$ . As it is usually required for trains arrive at stations within a minute range of proposed arrival time. In our study, 1% of total journey time is precise enough for this requirement.

The pre-determined position is the position at which the speed of the vehicle needs to be determined, shown in Figure. 5.1. Since the number of pre-determined positions for a known journey can be infinity, this complexity of search space will be not possible to implement into any optimisation algorithm. In order to address this issue, the pre-determined positions are firstly assigned for the journey and the searching algorithms are applied to search the speed for these pre-determined positions. The speed for the other positions is regarded as determined and uncontrollable. The calculation of speed at other positions is covered in Section 5.3.3. The pre-determined positions include the following three types, see Figure. 5.1 for details.

- Positions whose distance value is the multiple of the proposed distance interval  $s_{int}$ , e.g.  $S_2$  and  $S_3$  pre-determined positions in Fig. 5.1;
- Positions at which the speed limit is changed, e.g.  $S_5$  and  $S_7$  pre-determined positions in Fig. 5.1;
- Positions at the beginning and end of the journey, e.g.  $S_1$  and  $S_{11}$  pre-determined positions in Fig. 5.1.

For each pre-determined position, the possible speed should meet the constraints covered in Section 5.3.2 and 5.3.3.

#### 5.3.2 Speed limitation

The limitations of the traction system and the speed limits imposes restrictions on the speed which is achievable at various distances. To avoid any unfeasible speed,

it is necessary to run the train in a “Flat Out Running (FOR)” style to set the speed limit at each distance.

**Definition 5.3** (Flat Out Running). *Flat Out Running is such a train running style at which the train will run with its maximum possible speed at every position along a journey.*

FOR can be interpreted as follows:

1. When the train is accelerating, the speed at any position cannot be further increased due to the physical constraints and speed limit;
2. When the train is decelerating, the speed at any position cannot be further decreased or the speed will be decreased down to any lower level than a permitted speed;
3. When the train is cruising, the speed at any position will remain at the speed limit until acceleration or deceleration is activated.

Let  $v_{s_n}^{max}$  denote the maximum speed at each pre-determined position  $s_n$ , where the subscript  $n$  implies the index of the pre-determined position. For each pre-determined position, Eqn. 5.39 holds.

$$v_{s_n}^{max} \geq v_{s_n} \quad (5.39)$$

From a practical perspective, the train should not run at too low a speed level to avoid unpractical speed trajectory. Thus, a lower limit for the speed at each pre-determined position can be defined. The lower speed limitation can be defined by Eqns. 5.40 and 5.41.

$$v_{s_n}^{min} = C_p \cdot v_{s_n}^{max} \quad (5.40)$$

$$v_{s_n}^{min} \leq v_{s_n} \quad (5.41)$$

where  $C_p \in [0, 1]$ .

### 5.3.3 Minor speed switch

It is not possible for the train to accelerate or decelerate from any speed  $v_{cur}$  at the current pre-determined position to any speed  $v_{nxt}$  at the next pre-determined position even if both speeds are lower than the respective  $v_{s_n}^{max}$ , since the acceleration supplied by the traction system is limited by the current vehicle speed and motor characteristics.

The speed switch between two pre-determined positions is done through the minor simulation distance interval for a greater calculation precision. Within a minor simulation distance interval, the resistance and tractive force are regarded as constant and thus the acceleration is constant. Details of minor simulation can be found in Section 4.3.1.

Two speeds are regarded as connectable if the train is able to switch its current speed  $v_{cur}$  to  $v_{nxt}$  at the distance interval. The possible connection between two speeds at two different pre-determined positions will be first to evaluate. The model runs the train using the “FOR and brake as hard as possible running” style until the train arrives at the next position. The  $v_{fo}$  is used to denote current speed after the “FOR” and  $v_{bh}$  to denote the speed after the “brake as hard as possible”. If  $v_{fo} \geq v_2$  and  $v_{bh} \leq v_2$ , it is regarded that these two speeds are connectable for the current distance interval, otherwise it is regarded that they are not connectable. It is noted that the connection is only possible for the speed for any forward adjacent pre-determined positions as shown in Figure 5.3(a).

However, for two speeds at two pre-determined positions, the number of possible speed trajectories is infinite, as long as the speed changing rate, i.e. acceleration, does not exceed the instant limitation, see Figure 5.2. Two types of trajectory are defined in this model to represent a typical train trajectory: one is a coasting trajectory and the other is a normal trajectory. For the coasting trajectory, the train is applying no extra tractive force and no traction power is consumed. The speed  $v_c^{nxt}$  for the next pre-determined position due to coasting is calculated first. If any

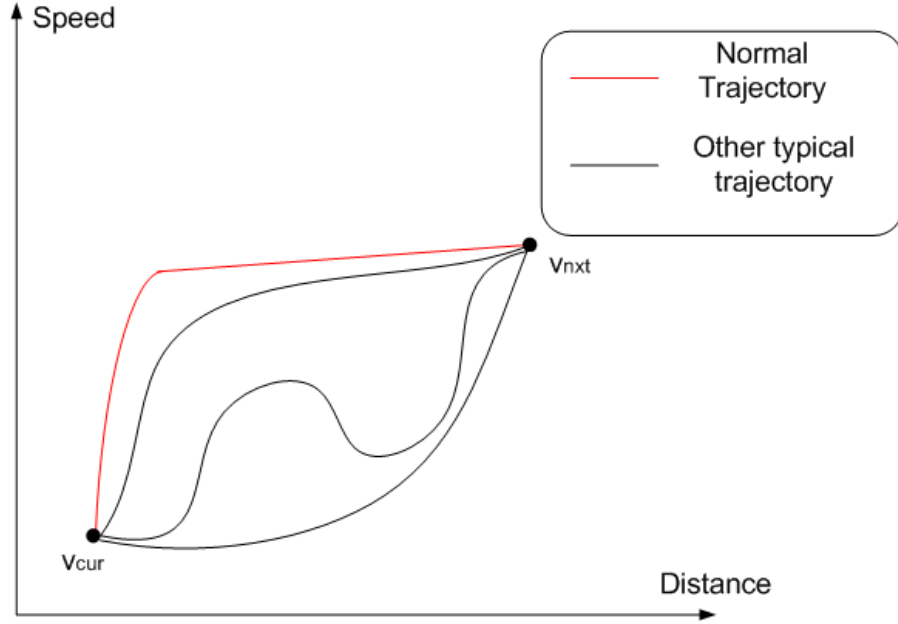


Figure 5.2: Various typical trajectories between two pre-determined positions

speed in the neighbourhood of current speed  $v_{cur}$  is equivalent to the  $v_c^{nxt}$ , the power consumption for the speed change will be set to zero, otherwise a normal trajectory will be adopted.

A normal trajectory is selected to represent general train running as this will reflect the most possible practical running of train in practice. This typical trajectory is obtained by a simple driving style that the train keeps to use the maximum available acceleration or deceleration defined by Figure 4.1 and Eqn. 4.9 until the speed of the train is close enough to  $v_{nex}$ . The vehicle then changes into constant small acceleration mode. The small acceleration is defined by Eqn. 5.42.

$$\tilde{a} = \frac{v_{nxt}^2 - v^2}{2 \cdot s_d} \quad (5.42)$$

where  $\tilde{a}$  is a small constant acceleration rate and  $s_d$  is the distance difference between train instant position and next pre-determined position,  $v$  is the instant speed.

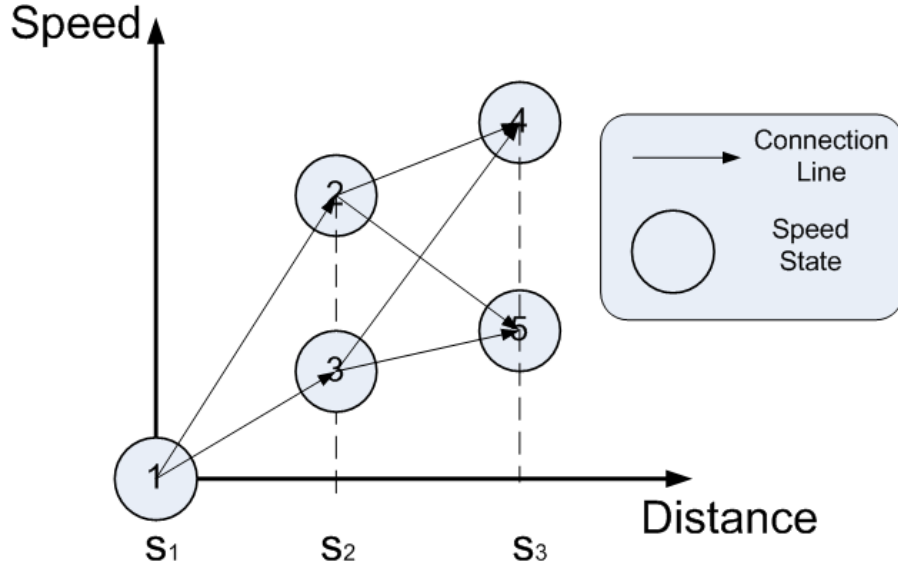
### 5.3.4 Sparse data storage

In order to enhance the searching speed of the algorithms, the time and energy consumption information for speed changes between two pre-determined positions are stored using a sparse matrix. For each feasible speed at each pre-determined position, a unique index number is allocated in the order of the journey forwarding direction. In this case, index number for state at the first pre-determined position will be higher than the one at the last pre-determined position. For any two connectable speeds  $v_i$  and  $v_j$ , the element at  $i$ th row and  $j$ th column of the sparse matrix is set to a nonzero number. The value of the nonzero number depends on the function of the sparse matrix. For example, if the matrix is for energy consumption, the data will be set to energy consumption for the speed change. The matrix can also be used to store the time consumption and the heuristic and linkage information for the speed change for the ACO algorithm, covered in detail in Section 5.4.

## 5.4 ACO application

### 5.4.1 Introduction

ACO is inspired by the foraging behaviour of the ant colony [84]. In ACO, a set of artificial ants communicate and cooperate indirectly by pheromone to find a solution to a difficult discrete optimisation problem. Each artificial ant, as an independent agent, is allocated with the computational resources by which it is able to leave the pheromone when necessary to communicate with the other ants. The ant with the good solution tends to leave more pheromone along its routes to direct the other ants. Using this learning enhancement style algorithm, the route with the better solution will finally attract more ants to follow and finally lead a convergence of the optimisation process.



(a) Speed link matrix diagram demonstration

| Speed Index | Speed 1 | Speed 2 | Speed 3 | Speed 4 | Speed 5 |
|-------------|---------|---------|---------|---------|---------|
| Speed 1     | —       | 1       | 1       | 0       | 0       |
| Speed 2     | 0       | —       | 0       | 1       | 1       |
| Speed 3     | 0       | 0       | —       | 1       | 1       |
| Speed 4     | 0       | 0       | 0       | —       | 0       |
| Speed 5     | 0       | 0       | 0       | 0       | —       |

(b) Speed link matrix table demonstration

Figure 5.3: Demonstration of the speed link matrix

### 5.4.2 Construction graph

As introduced in Section 5.3.4, the candidate speed is stored in sparse data storage and each candidate speed at each pre-determined position has its own unique index. A construction graph is a complete weighted graph  $G = (N, A)$  with  $N$  being the set of  $N$  nodes, i.e. the candidate speeds and “ $A$ ” being the set of arc connections between the nodes. Each  $i, j \in A$  is assigned with the consumed time  $TC(i, j)$  and energy  $EC(i, j)$ , assuming the train is running between two positions with the initial speed being  $v_i$  and final  $v_j$ .  $i, j \in A$  corresponds to the unique index for each node. As the train is not able to travel backwards and each speed has its own unique

index, it is obvious that  $TC(j, i)$  is zero if  $TC(i, j)$  is a real number.

Some definitions are made as follows:

- Let  $EC$  denote the sparse matrix to store the energy consumption of every two nodes. When the train is unable to change its speed from the initial speed of node  $i$  to the final speed of node  $j$  within the distance between the positions of the two speeds, the  $EC(i, j)$  is set zero. Otherwise a nonzero number will be set for  $EC(i, j)$  to represent the energy consumption when train is switching between these two states. Note that any negative energy consumption will not be added into the final energy consumption assuming that no power is consumed due to braking.
- Let  $TC$  denote the sparse matrix to store the time consumption of every two nodes. Any unconnected nodes are set as zero, and a nonzero number of  $TC(i, j)$  otherwise.
- Let  $ECH$  denote the sparse matrix to store the heuristic energy consumption of every two nodes. Any unconnected nodes are set as zero, and a nonzero number of  $|1/EC(i, j)|$  otherwise. Note that for zero energy consumption the element of  $ECH(i, j)$  is infinity. The heuristic consumption implies the magnitude of the energy saving. For example, the higher the heuristic energy consumption, the lower the energy cost will be and the higher the energy saving will be. The function of heuristic value will take the phenomenon effect to help the artificial ant to decide the next candidate speed, and this will be covered in Section 5.4.3.
- Let  $TCH$  denote the sparse matrix to store the heuristic time consumption of every two nodes. Any unconnected nodes, the element  $TCH(i, j)$  is set as zero, or a nonzero number, i.e.  $1/TC(i, j)$  otherwise. The heuristic consumption implies the significance of time saving, i.e. the higher the heuristic time consumption value is, the less time is consumed and less time is caused.

- Let  $LNK$  denote the pheromone information sparse matrix. Each connected arc in the weighted graph  $G=(N,A)$  can be selected due to the pheromone information originally imposed. The pheromone information originally imposed contains two parts:
  - The linkage information between two nodes. As the pheromone will not be reduced down to zero after several updates, the nonzero characteristic will be kept to indicate whether these two nodes can be connected.
  - The strength of linkage between two nodes. A higher value between two indexed nodes in  $LNK$  suggests a stronger linkage between two nodes. Hence it is more possible for an ant to do a speed change between these two linked nodes.

### 5.4.3 Solution construction

The pheromone trail  $LNK(i, j)$  refers to the desirability of selecting the next candidate speed with the index of “ $j$ ” from the current candidate speed with the index of “ $i$ ”. Because the possibility of connection between two candidate speeds only relies on the two speeds of the adjacent pre-determined positions, every selection step will push the train’s position forwards until the final station position.

The original pheromone trail imposed for every connectable two nodes is a constant  $c_{lnk}$  shown as follows.

$$LNK(i, j) = c_{lnk} \quad (5.43)$$

At each construction step, ant “ $k$ ” chooses the next speed at the next pre-determined position based on a random proportional rule [84, P. 70]. Assume that the ant is currently at the speed index  $i$  and the possibility of speed index  $j$  being selected is defined as follows:



$$P_{i,j}^k = \frac{[LNK(i, j)]^\alpha [ECH(i, j)]^\beta [TCH(i, j)]^\gamma}{\sum_{n \in \Omega_i^k} [LNK(i, n)]^\alpha [ECH(i, n)]^\beta [TCH(i, n)]^\gamma} \quad (5.44)$$

where  $LNK(i, j)$ ,  $ECH(i, j)$  and  $TCH(i, j)$  are the linkage pheromone trail, energy heuristic consumption and time heuristic consumption respectively.  $\alpha, \beta, \gamma$  are three parameters to determine the relative influence of the pheromone trail and the heuristic information, and  $\Omega_i^k$  are the feasible neighbourhood of ant “k” being at node i. In the sparse matrix data storage, the neighbourhood of the node “i” is defined by an index of nonzero elements of the “ith” row of the linkage matrix. Take Figure 5.3(b) as an example, the row of speed 1 has two nonzero elements for the linkage matrix, i.e. speed 2 and speed 3. These two speeds are in the neighbourhood  $\Omega_1$  of speed 1 (Node 1).

As it is not straightforward to decide which heuristic information should be taken and it is not possible to obtain fast convergence if both are undefined [84], some other information should be considered. The following concepts are defined:

- **Average speed** The average speed is defined by the proposed journey time and total journey length.

$$v_a = \frac{S_t}{T_{prop}} \quad (5.45)$$

where  $S_t$  is the journey length,  $T_{prop}$  is the proposed journey time,  $v_a$  is the average speed.

- **Proposed arrival time** The proposed arrival time for the “Mth” pre-determined position can be defined as follows:

$$t_{prop}^M = \frac{S_i}{v_a} \quad (5.46)$$

- **Actual arrival time** The actual arrival time for the “Nth” pre-determined

position can be defined as follows:

$$t_{act}^N = \sum_{n=1}^{N-1} tc_{n,n+1} \quad (5.47)$$

where  $tc_{n,n+1}$  is the time consumption between “nth” pre-determined position and “(n+1)th” pre-determined position. Assume that for the “nth” pre-determined position the selected speed has the index “i” and for “(n+1)th” pre-determined position the selected speed is with the index “j”, the  $tc_{n,n+1}$  can be obtained from the following equation.

$$tc_{n,n+1} = TC(i, j) \quad (5.48)$$

where  $TC$  is the sparse storage for time consumption where i and j are the indexes for the two candidate speeds at two pre-determined position.

- **Time margin** The time margin at the “Nth” pre-determined position is defined as follows:

$$t_{mar}^N = t_{act}^N - t_{prop}^N \quad (5.49)$$

Intuitively, if  $t_{mar}^N$  is positively larger, it generally means that the train needs to keep up its speed to catch up with the average speed.

A time constant  $T_c$  is defined to evaluate the current state of train running in terms of punctuality and  $p_u = t_{act}/T_{prop}$  as a ratio of the current used time over total proposed journey time. According to the above-mentioned definitions, by evaluating the time margin, the dynamic changing heuristic value can be determined by simple rules listed in Table 5.1. The values listed in Table 5.1 are selected through a “trial and error” means to realise an effective search performance.

Based on the rules listed in Table 5.1 and by applying Eqn. 5.44, each artificial ant is able to specify the next indexed speed for the next pre-determined position and finally a resultant journey can be constructed showing the speed of the train at

## 5.4. ACO APPLICATION

Table 5.1: Rules to decide heuristic coefficients

| Conditions                         | Heuristic Value   | Notes   |
|------------------------------------|---|---|
| $t_{mar} \geq 0$                   | $\alpha = 1$<br>$\beta = 0$<br>$\gamma = 3^{(p_u)}$           | The train is not quick enough<br>the time becomes more important<br>with the used time approaching the end                                  |
| $0 > t_{mar} \geq -T_c$            | $\alpha = 1$<br>$\beta = 0$<br>$\gamma = 1$                   | The train is fast enough.<br>Small emphasis with<br>the parameter $\gamma = 1$ is imposed.  |
| $-T_c > t_{mar} \geq -2 \cdot T_c$ | $\alpha = 1$<br>$\beta = 3$<br>$\gamma = 0$                   | Train is fast enough to<br>impose the energy efficiency operation<br>Time is not considered   |
| $-2 \cdot T_c > t_{mar}$           | $\alpha = 1$<br>$\beta = 5$ if $p_u \leq 0.8$<br>$\gamma = 0$ | If time used is near the end, the next<br>speed will be determined purely by<br>the pheromone trail, setting $\beta = 0$ in-<br>stead of 5. |

each of the pre-determined positions. Using the pre-acquired two sparse matrixes  $TC$  and  $EC$ , the journey time and energy consumption can be obtained. The quality of the solution will be evaluated using the evaluation function shown in Eqn. 5.36 and the pheromone trail will then be updated based on the quality of each constructed solution. One of the key functions of the update procedure is to reinforce the better solution by imposing more pheromone trail. Further details are covered in Section 5.4.4.

### 5.4.4 Pheromone update and termination condition

The pheromone trail matrix is updated using the output of Eqn. 5.44. In this chapter, a generalised update procedure is adopted for a group of artificial ants.

Use  $n_a$  to denote the number of ants in a group. Let  $n_p$  be the number of pre-determined positions. Use  $SOL$  to denote the constructed solutions matrix, with each row being a constructed solution. For each solution, the element is the index of each candidate speed at each of the pre-determined positions. The number of elements in each row equals to  $n_p$ . Use  $EVAL$  to denote the one dimension matrix

---

#### 5.4. ACO APPLICATION

---

to store the evaluation function output for each row of constructed solutions. Let  $UPD$  denote the update vector to update the pheromone trail.

The update procedure can be divided into two parts. The first part is illustrated in the pseudo-code shown in Algorithm 5.1.

---

**Algorithm 5.1** Part I: obtain the update vector  $UPD$  for each constructed solution in  $SOL$

---

**Require:**  $eval_{min} \leftarrow \min(EVAL)$   
**for**  $i = 1$  to  $n_a$  **do**  
     $eval = EVAL(i) - eval_{min}$   
     $UPD(i) = 2 \cdot c_{lnk} \cdot \exp(-eval)$   
**end for**

---

Please note that  $\min$  is a function which is to obtain the minimum element from its input vector,  $\exp$  is the exponential function and  $c_{lnk}$  is defined in Eqn. 5.43. The second part is illustrated in the pseudo-code shown in Algorithm 5.2.

---

**Algorithm 5.2** Part II: update the pheromone trail matrix  $LNK$  using  $UPD$  and  $SOL$

---

**for**  $i = 1$  to  $n_a$  **do**  
    **for**  $j = 1$  to  $n_p - 1$  **do**  
         $r_i \leftarrow SOL(i, j)$   
         $c_i \leftarrow SOL(i, j + 1)$   
         $LNK(r_i, c_i) \leftarrow (1 - c_e)LNK(r_i, c_i) + UPD(i)$   
    **end for**  
**end for**

---

The best solution searched so far  $sol_{bsf}$  will be stored and updated by the new solution if a lower evaluation function output can be achieved.

The termination condition is set as two criteria. Firstly, the number of groups of ants exceeds the maximum allowable number. Secondly, the  $sol_{bsf}$  keeps unchanged for a selected number of iterations.

## 5.5 Genetic algorithm

### 5.5.1 Introduction

The GA as a population based optimisation does not require gradient information of the objective function and only uses the output of the function to guide the search for a better solution. Due to this fact, the heuristic information in a GA is not as important as that in ACO where the search for the solution is well guided by current heuristic information, i.e. is the train fast enough to coast. As mentioned in Section 5.1, the GA has been reported as the successful candidate algorithm in various train speed trajectory searching applications and the simulation results show its robustness in this area [25,26,33,133]. In this section, a GA will be used to search for the characterised speed at each pre-determined position using modelled strings (genotypes). Each string is modelled as a characterised signal for the current speed jump. The GA computational tool by MATHWORKS is used [138]. Further details are covered in Section 5.5.2. A similar application of GA can be found in [33], where the control signals at each position are modelled into one matrix. The number of positions is allowed to change and a multiple population GA is applied to search for a good control signal sequence for each of the corresponding positions.

Genetic algorithm takes its significance in optimisation without requiring gradient information of the evaluation function in relation to the input variable. Consequently, it is suitable in searching for the minimum or maximum of a nonlinear function, which cannot be analytically presented. Differing from ACO, GA does not rely on heuristic information which can guide the search procedure into a relatively fast convergence. The guidance is taken from the output of strings (genotypes) in one generation and GA makes sure that better strings have more chance of affecting the future search direction. This is why the result achieved by GA may vary within a larger range in comparison to ACO. This observation is covered in Section 5.7

### 5.5.2 Genotype generation

In order to apply the GA, two important steps should be implemented before different operators of GA can apply.

- Generation of the population of strings (genotype).
- A fitness function should be created to distinguish the better strings from the poorer ones.

As all of the strings should be able to gain the final journey time and total energy consumption, the objective function Eqn. 5.36 used in Section 5.4 can be still used. The first step to generate feasible strings is important as all these strings deliver an applicable train trajectory.

Due to the non-linearity between the adjacent two speeds, it is not applicable to model speed at each pre-determined position into one candidate string. However, note that for each candidate speed at each pre-determined position, the speeds in the neighbourhood do have a range. Referring back to Section 5.2, various characterised operations can be identified through the speed change. For example, characterised full motoring can be referred to as the speed change by selecting the fastest speed in the neighbourhood and characterised full braking can be referred to as changing the speed by selecting the slowest speed in the neighbourhood.

Figure 5.4 shows the speed change can be achieved by using various characterised control signals for each pre-determined position interval. The different Arabic numbers indicate different types of characterised control signals. Let  $i_c$  denote the control index number for the characterised control. Let  $V_{next}^{max}$  and  $V_{next}^{min}$  denote the next possible maximum speed and minimum speed in the neighbourhood of the current speed  $V_{cur}$  in Figure 5.4.  $v_{ee}$  the next speed through the most energy efficient operation is defined as follows:

**Definition 5.4** (Next speed through the most energy efficiency operation). *If a speed  $v_j$  falls into the neighbourhood of  $v_i$  with “ $i$ ” and “ $j$ ” as the unique index,*

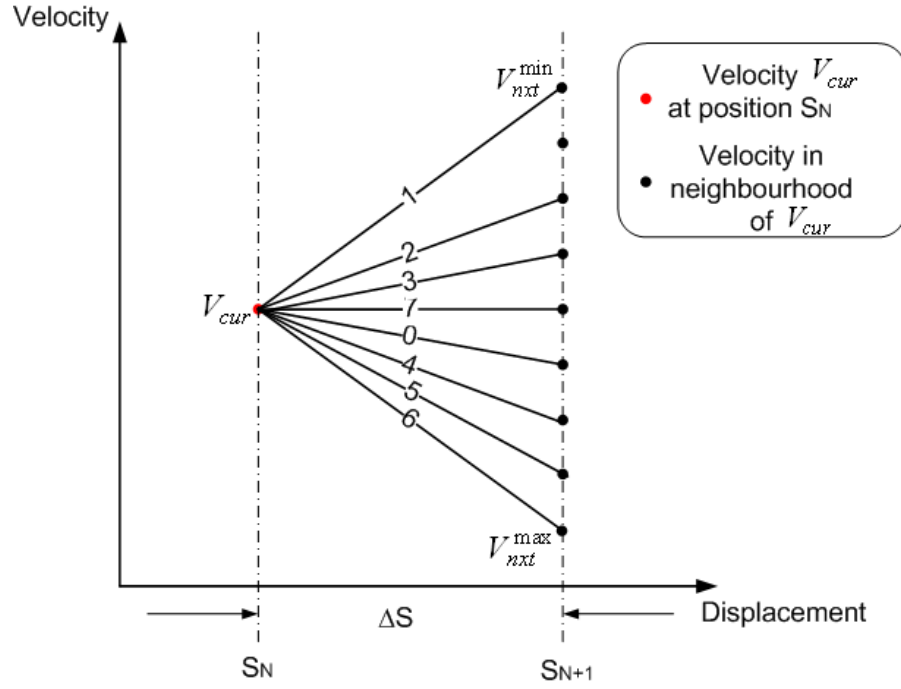


Figure 5.4: Various speed change signal for one pre-determined position interval

Table 5.2: Next speed selection based on the characterised control index number

| Control index $i_c$ | Next speed selected                                     |
|---------------------|---|
| 0                   | $v_{ee}$ or $v_c$ whichever exists                      |
| 1 ~ 6               | $V_{min} + (V_{max} - V_{min}) \cdot \frac{i_c - 1}{5}$ |
| 7                   | $V_{cur}$ if feasible                                   |

and the following relation holds:

$$LNK(i, j) \geq LNK(i, n) \quad \forall n \in \omega_i \quad (5.50)$$

where  $\omega_i$  is the neighbourhood of the speed  $v_i$ .  $v_j$  is defined as the next speed through the most energy efficiency operation.

Let  $v_{ee}$  denote such  $v_j$  for any speed  $v_i$  if  $LNK(i, j) \neq \infty$  and let  $v_c$  denote such  $v_j$  if  $LNK(i, j) = \infty$ .

Since heuristic energy consumption is the absolute reciprocal of real energy consumption, the maximum heuristic energy consumption indicates the minimum absolute real energy consumption of the train. Though negative energy consump-

tion is not counted in the total energy consumption, the kinetic energy converted from the positive power will still be consumed as heat. From the point of view of energy conservation, the absolute energy consumption is used to evaluate the energy efficiency of each speed change. If the coasting operation is selected, i.e.  $v_c$  exists,  $v_c$  will be selected as the speed of the next pre-determined position.

For the control index number of “0”, the most energy saving mode is selected. For the control index of “1” to “6”, six speeds in the range of  $[V_{nxt}^{min}, V_{nxt}^{max}]$  will be selected using the equation shown in Table 5.5.2. The control index number of “7” implies a cruising operation for the vehicle if the speed is allowed to be maintained in the next pre-determined position.

In view of the above illustration, if a string contains the control index number for each pre-determined position interval, the speed can be easily determined at each pre-determined position given that the first speed at the first pre-determined position is zero. Using the pre-acquired matrices  $TC$  and  $EC$  and the unique speed index, the energy and time consumption can be obtained. Finally, the fitness of each string can be evaluated and guide the GA operation to achieve a better trajectory.

Assume that there are “M” strings in the initial population and “N” pre-determined positions along the journey, the schematic GA optimisation procedure is shown in Figure 5.5

The details of the GA operation is covered are Section 3.4.2.

## 5.6 Dynamic programming

### 5.6.1 Introduction

Dynamic programming is a powerful tool which can be used to solve a problem which can be divided into various sub-stages. In this case, trajectory search can naturally be divided into sub-intervals of the distance.



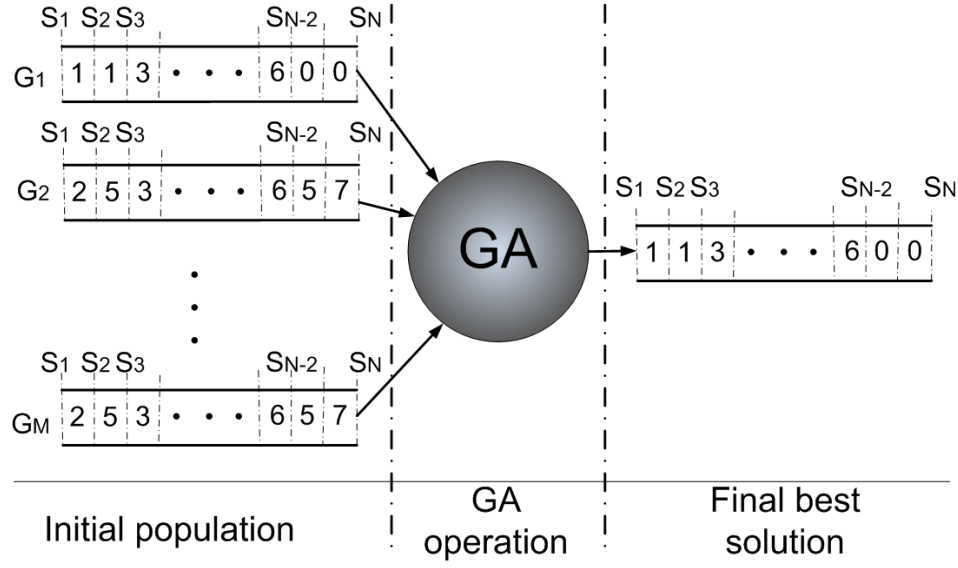


Figure 5.5: Schematic Genetic Algorithm optimisation of train speed trajectory.

### 5.6.2 Optimisation process

Firstly the following two definitions are made:

**Definition 5.5** (Vehicle state). *Vehicle state  $\phi$  consists of three basic physical elements: vehicle distance  $s$ , vehicle speed  $v$  and used journey time  $t$  since setting off from the initial state where  $s = 0$  and  $v = 0$ .*

*Vehicle state can be expressed in an array form*

$$\phi = [s, v, t] \quad (5.51)$$

**Definition 5.6** (Augmented vehicle state). *Augmented vehicle state  $\phi^a$  contains other elements which mainly serve for the purpose of search except the elements in  $\phi$ .  $\phi^a$  contains vehicle distance  $s$ , vehicle speed  $v$ , instant used journey time  $t$ , the unique speed index for instant vehicle speed  $v$ <sup>1</sup> and the unique index for current vehicle state from which current augmented vehicle state is produced and the minimum energy cost resulted from initial vehicle state  $\phi_o$ . For more information on unique speed index, see Section 5.3.4.*

---

<sup>1</sup>Any group of speed and distance will define a unique index in the sparse matrix.

$\phi^a$  can be expressed in a array format as:

$$\phi^a = [s, v, t, i, \hat{i}, e] \quad (5.52)$$

Note that for each augmented vehicle state  $\phi^a$ , there is one corresponding vehicle state  $\phi$ .

$$\phi_o^a = [s_o, v_o, t_o, i_o, \hat{i}_o, e_o] = [0, 0, 0, 1, 0, 0] \quad (5.53)$$

By the presumption, each vehicle state must have one of the pre-determined positions as its current distance information.

The following definitions are made:

- Let  $U$  define the control signal vector for the vehicle state and  $u_x$  denote one of the control signal numbers.  $u_x = 0, 1, 2 \dots 7$
- Let  $S_n$  denote the “ $n_{th}$ ” pre-determined position.
- Let  $\Phi_T$  denote the total state set which contains all the produced augmented vehicle states. In  $\Phi_T$ , the unique index  $i$  and  $\hat{i}$  for each state will help locate the state itself and the previous state.
- Let  $Ind(s, v)$  denote the speed index in the sparse storage matrix for each speed, see Section 5.3.4 for more detail.
- Let  $\Phi_n^a$  denote the state set including all the created augmented vehicle states at the pre-determined position  $S_n$ . Note that state set  $\Phi_{n+1}^a$  is produced from the set  $\Phi_n^a$  using control signals  $U$  shown in Eqn. 5.54.

$$\Phi_n^a \xRightarrow{U} \Phi_{n+1}^a \quad (5.54)$$

The initial state set  $\Phi_1^a$  is initialised using Eqn. 5.55.

$$\Phi_1^a \Leftarrow \phi_o^a \quad (5.55)$$


---

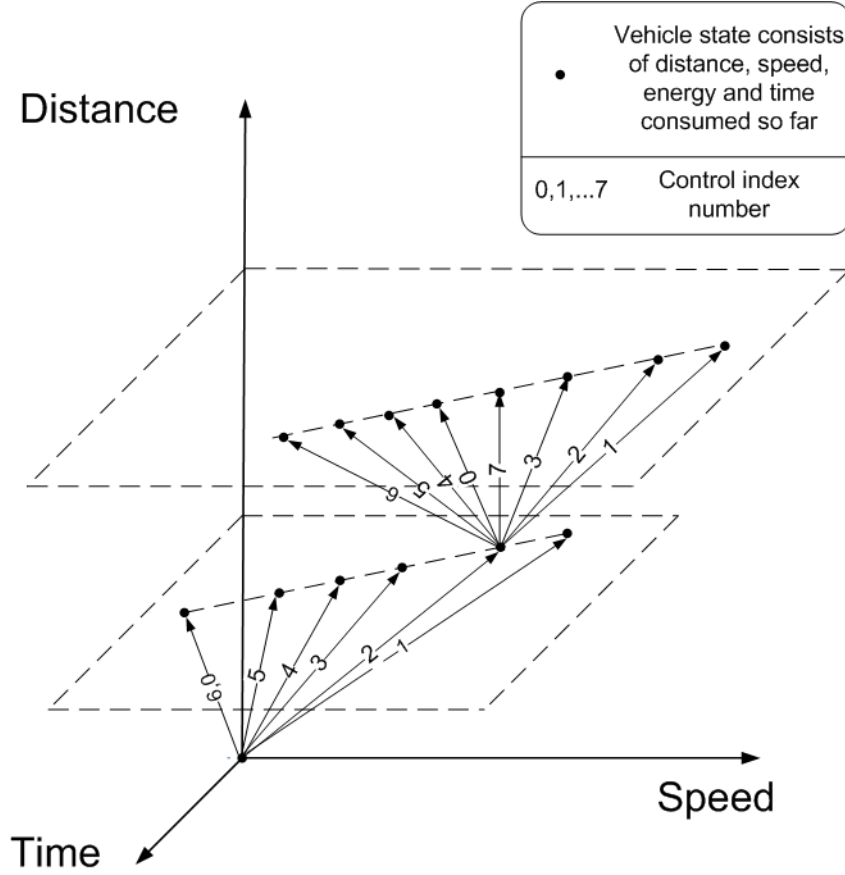


Figure 5.6: Vehicle states generation procedure in DP.

“ $\Rightarrow$ ” denotes the procedure where the vehicle state is switched through a physical change, e.g. distance, and a speed change will result in an energy and time cost.

“ $\Leftarrow$ ” denotes the storing procedure through which the vehicle state is stored in a matrix storage format and each different state is indexed with a unique index number for further search purposes.

“ $\rightarrow$ ” denotes the new state’s creation procedure from a single augmented vehicle state by applying all of the control input.

- Let  $\phi_n[m]$  denote one of the various vehicle states and  $\phi_n^a[m]$  denote one various augmented vehicle state at the  $n_{th}$  pre-determined position.

The Dynamic Programming proceeds forwards iteratively as shown in Figure 5.6. In Figure 5.6, all the vehicle states are developed from the initial one, which

is the original point in this graph. For the first step, the states will be created from the initial state using the index control signal mentioned in Section 5.5. The index control signal for a cruising operation is not available for the first step. One of the examples for the second step is also presented which should be applied for all the other created states from the initial state. Further details of this procedures are described as follows.

1. Assume that there are  $N$  pre-determined positions along the journey. For the first pre-determined position, the vehicle state is  $\phi_o$ . Set the current iteration index  $n = 1$  and store the initial augmented vehicle state into the  $\Phi_1^a$  as in Eqn. 5.55.
2. For all the elements of  $\Phi_1$ , i.e. only one element in  $\phi_o$ , in the neighbourhood of  $v_o$ , apply all the control signal number  $u_x$  for  $v_o$  and obtain the next speed  $v_x$ . Obtain the unique index  $i_x$  of each speed and the energy cost  $e_x$  since the initial state  $\phi_o$  respectively. Obtain the new vehicle states based on the speed change using the following equations.

$$s_x = S_2 \quad (5.56)$$

$$t_x = TC[Ind(s_o, v_o), Ind(s_x, v_x)] \quad (5.57)$$

$$\hat{i}_x = i_o = 1 \quad (5.58)$$

$$e_x = EC[Ind(s_o, v_o), Ind(s_x, v_x)] \quad (5.59)$$

where TC and EC are time cost and energy cost sparse matrix respectively as defined in Section 5.3.4, through which the energy and time cost for the vehicle switch state from  $\phi_o$  to  $\phi_x$  can be obtained.  $\phi_x^a$  is used to denote any of the new vehicle augmented states, defined as:

$$\phi_x^a = [s_x, v_x, t_x, i_x, i_o, e_x] \quad (5.60)$$

The new vehicle augmented states creation procedure can be presented as

shown in Eqn. 5.61:

$$\phi_o^a \xrightarrow{u_x} \phi_x^a \quad (5.61)$$

The distance in  $\phi_x^a$  will be the next pre-determined position. The speed in the new vehicle augmented state must be in the neighbourhood of the speed in the previous state. The time used must be increasing because of the positive time cost for the vehicle state switch.

3. As for the initial augmented vehicle states  $\phi_o^a$ , a set of augmented vehicle states  $\phi_x^a$  can be produced and stored into the storage  $\Phi_2^a$ . This procedure will not end until all the elements of the previous vehicle state set  $\Phi_1^a$  have been used to create the new vehicle augmented states.

If there are any two augmented states  $\phi_2^a[x_1]$  and  $\phi_2^a[x_2]$  with the identical vehicle state, the energy consumed from the initial state to the current state in the two augmented states will be compared. The augmented vehicle state with the lower energy cost will remain and the other one will be eliminated. This procedure will be performed iteratively. As a special case, if the energy consumption is the same, either of the augmented states will be selected. For example, there are two new augmented vehicle states shown as follows.

$$\phi_2^a[x_1] = [s_{x_1}, v_{x_1}, t_{x_1}, i_{x_1}, \hat{i}_{x_1}, e_{x_1}] \quad (5.62)$$

$$\phi_2^a[x_2] = [s_{x_2}, v_{x_2}, t_{x_2}, i_{x_2}, \hat{i}_{x_2}, e_{x_2}] \quad (5.63)$$

These two augmented vehicle states have the same vehicle state, which means the following relationship holds.

$$s_{x_1} = s_{x_2} \quad (5.64)$$

$$v_{x_1} = v_{x_2} \quad (5.65)$$

$$t_{x_1} = t_{x_2} \quad (5.66)$$

If  $e_{x_1} < e_{x_2}$ ,  $\phi_{x_1}^a$  is a better route for the vehicle switch from initial state  $\phi_o$  to  $\phi_{x_1}$ ,  $\phi_{x_2}^a$  will be eliminated as a result.

By comparing every two augmented vehicle states with the same vehicle states, by deleting the state with a higher energy cost, each existing augmented vehicle state  $\phi_2^a[x]$  will finally contain the minimum energy cost resulting from the initial state  $\phi_1[1]$  to  $\phi_2[x]$ . By doing this recursively, at the final pre-determined position, all the augmented vehicle states in  $\Phi_N$  will contain the minimum energy cost results from the initial state.

As mentioned before,  $e_{n,m}$  denotes the minimum energy cost which results from the initial vehicle state  $\phi_1[1]$  to  $\phi_n[m]$  which denotes one of the state at pre-determined position  $S_n$ . Also, assume that at the  $n_{th}$  pre-determined position, there are  $M_n$  various vehicle states.

For the vehicle states at  $S_2$ , there is only one initial state at the “1st” pre-determined position  $S_1$ .

$$e_{2,1} = e_{1,1} + EC[Ind(s_{1,1}, v_{1,1}), Ind(s_{2,1}, v_{2,1})] \quad (5.67)$$

$$e_{2,1} = e_{1,1} + EC[Ind(s_{1,1}, v_{1,1}), Ind(s_{2,2}, v_{2,2})] \quad (5.68)$$

...

$$e_{2,m} = e_{1,1} + EC[Ind(s_{1,1}, v_{1,1}), Ind(s_{2,m}, v_{2,m})] \quad (5.69)$$

...

$$e_{2,M_2} = e_{1,1} + EC[Ind(s_{1,1}, v_{1,1}), Ind(s_{2,M_2}, v_{2,M_2})] \quad (5.70)$$

For a more general case, for any  $e_{n,m}$  where  $n = 1, 2, 3, \dots, N$  and  $m = 1, 2, 3, \dots, M_i$ . The minimum energy cost states route for a vehicle augmented state  $\phi_n^a[m]$  can be derived from the minimum energy cost of any possible previous augmented state  $\phi_{n-1}^a[m_{n-1}]$  with its minimum energy cost since initial state  $\phi_1^a[1]$ . Please note that possible previous augmented state  $\phi_{n-1}^a[m_{n-1}]$ , means that there should be at least one control signal number  $u_x$  leading

---

$\phi_{n-1}[m_{n-1}]$  to  $\phi_n[m_n]$ .

$$\phi_{n-1}[m_{n-1}] \xRightarrow{u_x} \phi_n[m_n] \quad (5.71)$$

where  $m_{n-1}$  and  $m_n$  are integers and the following relationships holds.

$$m_{n-1} \in [1, M_{n-1}] \quad (5.72)$$

$$m_n \in [1, M_n] \quad (5.73)$$

Assuming that there are  $m_p$  possible previous vehicle states,

$$\begin{aligned} e_{n,m} = & \min(e_{n-1,1} + EC[Ind(s_{n,m}, v_{n,m}), Ind(s_{n-1,1}, v_{n-1,1})], \dots, \\ & e_{n-1,m_p} + EC[Ind(s_{n,m}, v_{n,m}), Ind(s_{n-1,m_p}, v_{n-1,m_p})]) \end{aligned} \quad (5.74)$$

As a result, the minimum energy cost can be expressed in a general form.

$$e_{n,m} = \begin{cases} e_{1,1} + EC[Ind(s_{1,1}, v_{1,1}), Ind(s_{2,m}, v_{2,m})] & \text{for } n = 1 \\ \min(\varphi_{n-1}[1] + EC[Ind(s_{n,m}, v_{n,m}), Ind(s_{n-1,1}, v_{n-1,1})], \dots, \\ \varphi_{n-1}[m_p] + EC[Ind(s_{n,m}, v_{n,m}), Ind(s_{n-1,m_p}, v_{n-1,m_p})]) & \text{for } n = 2, \dots, N \end{cases} \quad (5.75)$$

4. Iteratively, from the initial state, the vehicle is able to switch to the other using Eqn. 5.54. As long as step 3 is performed whenever a duplicate vehicle is created, each state created will have the minimum energy cost since the generation of the initial vehicle state. Assume  $\phi_n^a[m]$  is one of the vehicle augmented states at pre-determined position  $S_n$ , defined as follows:

$$\phi_n^a[m] = [s_{n,m}, v_{n,m}, t_{n,m}, i_{n,m}, \hat{i}_{n,m}, e_{n,m}] \quad (5.76)$$

Use  $\phi_{n+1}^a[x]$  is one of the generated augmented states at pre-determined position  $S_{n+1}$ .

$$\phi_{n+1}^a[x] = [s_{n+1,x}, v_{n+1,x}, t_{n+1,x}, i_{n+1,x}, \hat{i}_{n+1,x}, e_{n+1,x}] \quad (5.77)$$


---

The following relation holds:

$$s_{n+1,x} = S_{n+1} \quad (5.78)$$

$$t_{n+1,x} = t_{n,m} + TC[Ind(s_{n,m}, v_{n,m}), Ind(s_{n+1,x}, v_{n+1,x})] \quad (5.79)$$

$$e_{n+1,x} = e_{n,m} + EC[Ind(s_{n,m}, v_{n,m}), Ind(s_{n+1,x}, v_{n+1,x})] \quad (5.80)$$

$$\hat{i}_{n+1,x} = i_{n+1,x} \quad (5.81)$$

where TC and EC are the sparse matrix of the time cost and energy cost for the speed switch defined in Section 5.4.2.

5. For the states in state set  $\Phi_N^a$  for the final pre-determined position, further selection should be performed to choose the best final state to meet the journey time requirement. It is noted that for all states in  $\Phi_N^a$ , the distance is the journey length and speed should be zero, while the time used varies. To meet the time requirement of the journey, the objective function is adopted as defined in Eqn. 5.36, to define the quality of the journey defined by final states in  $\Phi_N^a$ . As more emphasis is laid on the journey time, the “Penalty”  $P$  in Eqn. 5.36 is set to be as large enough to rule out any journey with a journey time cost more than the assigned upper bound  $b_u$  and lower bound  $b_l$ . The final state with the minimum energy cost which also meets the journey time requirement will be selected as the final augmented vehicle state.
6. After the final augmented vehicle state, denoted as  $\phi_f^a$ , is achieved, the total journey states route can be obtained through each unique index number for the previous vehicle augmented state.

To reduce the actual number of created states in the search space, a further heuristic action is taken to confine the actual search space. In [31], the “lattice area concept”, where the vehicle state can exist, is proposed. It is claimed that there are time-dependent admissible areas, i.e. lattice areas, from which the optimal train speed trajectory can be obtained. A vehicle state which is outside of the lattice

---



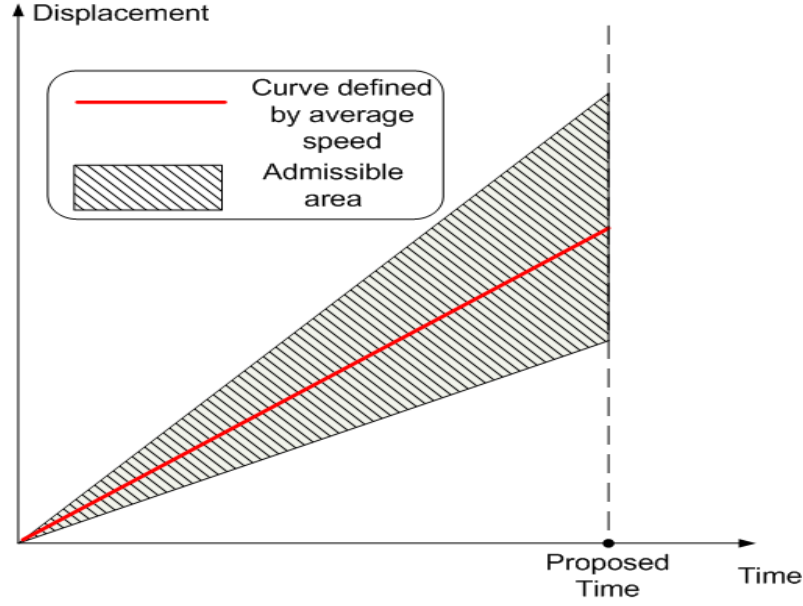


Figure 5.7: Admissible area of train states in DP search procedure

area will be ruled out from the search even if it is physically feasible. However, the reason why the area outside of the lattice area is regarded as qualitative non-optimal has not been proved. In our case, a simple heuristic is adopted to reduce search space and improve search efficiency: “It is generally accepted for a certain journey time cost, the instant position of the train should not vary too much from the position defined by the average speed.” This will significantly improve the search efficiency without sacrifice too much on the optimality of results. Accordingly, the upper bound and lower bound are defined for the journey time cost at different journey distances. As a result, an admissible area can be created. Figure 5.7 shows a similar concept for the admissible area in DP search.

### 5.6.3 Summary

In this section, DP has been applied to search for the optimum journey trajectories in terms of augmented vehicle state routes. By dividing the search procedure into sub-intervals, DP is able to obtain the minimum energy cost for a vehicle to switch from its original state to a current state by eliminating the same states with more energy cost. An admissible area for the DP search has been adopted to reduce the

## 5.7. RESULTS AND DISCUSSION

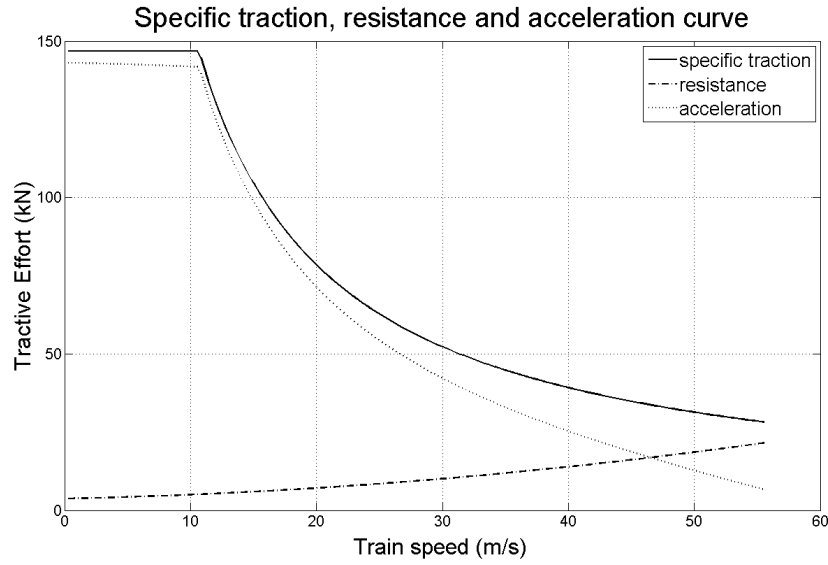


Figure 5.8: Maximum tractive effort, resistance and acceleration curve of Voyager type vehicle

Table 5.3: Key parameters for Single Train Motion Simulator

| Tare mass<br>(tonnes) | Maximum power<br>(kW) | Constant<br>torque (kN) | Davis coefficients |                                      |  |
|-----------------------|-----------------------|-------------------------|--------------------|--------------------------------------|--|
|                       |                       |                         | A (kN)             | B ( $\frac{\text{kN}}{\text{m/s}}$ ) | C ( $\frac{\text{kN}}{(\text{m/s})^2}$ ) |
| 213.19                | 1568                  | 146.8                   | 3.73               | 0.0829                               | 0.0043                                   |

total search states. Any states which stand outside of the admissible area will be ruled out of the searching procedure. The concept of admissible area relies on the heuristic that for a feasible solution of train trajectory, the instant position of the train cannot vary too much from the position defined by the average speed.

## 5.7 Results and discussion

The maximum tractive effort, resistance and corresponding maximum acceleration curve of the “Voyager” Class 220 is shown in Fig. 5.8.

Some of the key parameters for the Single Train Simulator as introduced in Section 4.3.4 are shown in Table 5.3.

---

## 5.7. RESULTS AND DISCUSSION

---

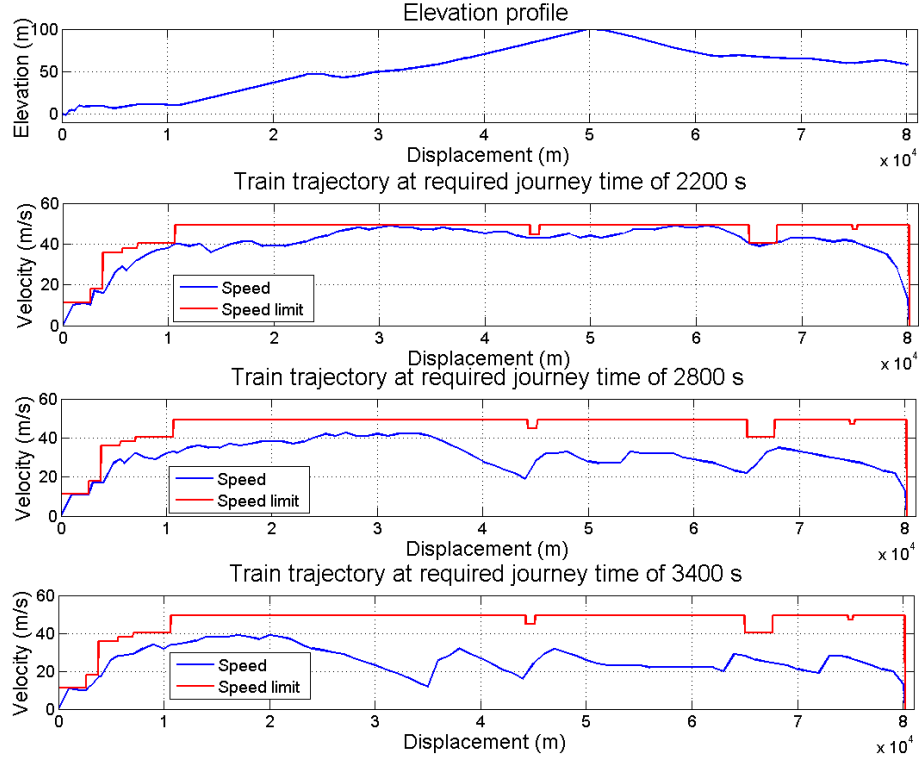


Figure 5.9: Optimised journey trajectories using ACO under different journey time conditions

Some of the simulation results are shown in this section. Firstly, trajectories searched under various journey time requirements, i.e. 2200 s, 2800 s, 3400 s for ACO, GA and DP are presented in Figure 5.9, Figure 5.10 and Figure 5.11.

The altitude profile of the journey is shown in each graph. With the increase in journey time, the speed trajectory is reduced to a lower level. Meanwhile, coasting operations (the zigzag part in the trajectory) become more preferable with a more journey time requirement. A comparison of the three trajectories shows that, using DP, the coasting operation is more frequently applied. With ACO or GA, however, the fluctuation of the speed is not as considerable.

Figure 5.12 and Figure 5.13 demonstrate the procedure in which objective function output evolves with the generation at the journey time requirement of 2800 s.

The journey time cost vs. energy cost curves for different journey time requirements are compared between the three algorithms. The journey time cost ranges from 2100 s to 3500 s with a 100-second increase interval. These curves can be

---

## 5.7. RESULTS AND DISCUSSION

---

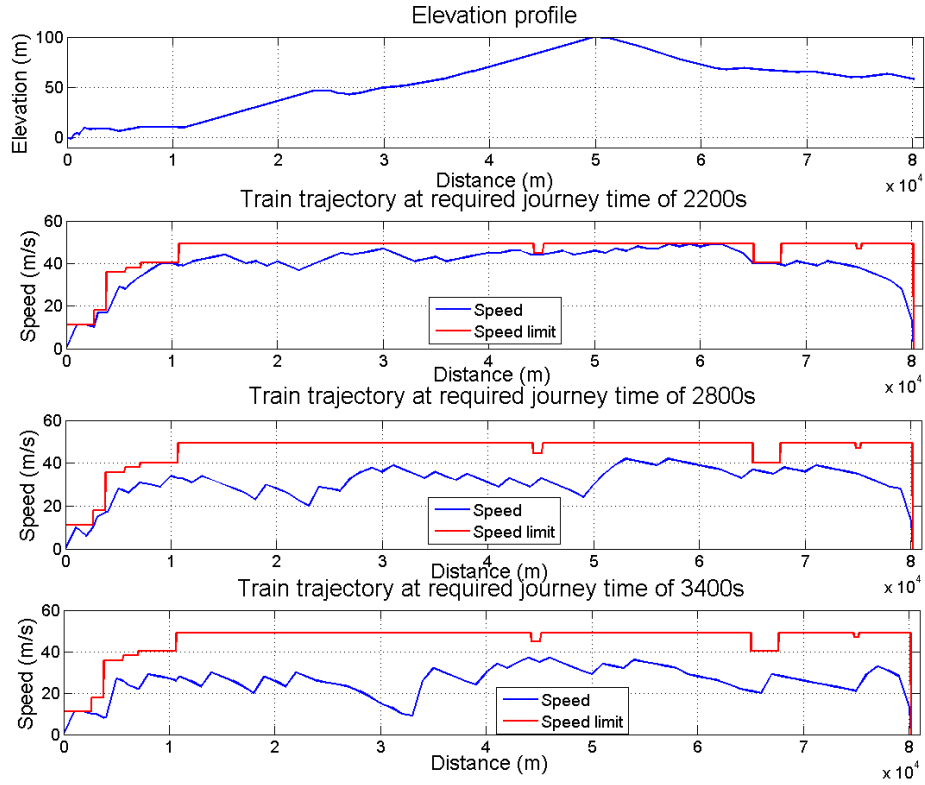


Figure 5.10: Optimised journey trajectories using GA under different journey time conditions

found in Figure 5.14. Both DP and ACO show their optimised journey trajectory using the cruising operation, during which the train is kept at a constant speed. Cruising tends to occur during the downhill part of the journey to take the advantage of the downhill gradients while maintaining reasonable average speed.

Note that each mark in the figure shows a combination of the journey time cost and energy cost for a simulation, and the colour and shape of the marks distinguish the type of algorithm. It is observed that the journey time cost for each algorithm is relatively consistent since the optimised journey time which is obtained is as close as possible to the required one. An interesting observation can be made that, while the GA maintains a significantly better performance in terms of a lower energy cost in comparison with ACO for the journey time requirement ranging from 2200 s to 3000 s, it becomes unstable and achieves a higher energy cost when the journey time is more than 3000 s. Since less heuristic information is adopted in the GA, the randomness affects the search result when the searching space expands with

## 5.7. RESULTS AND DISCUSSION

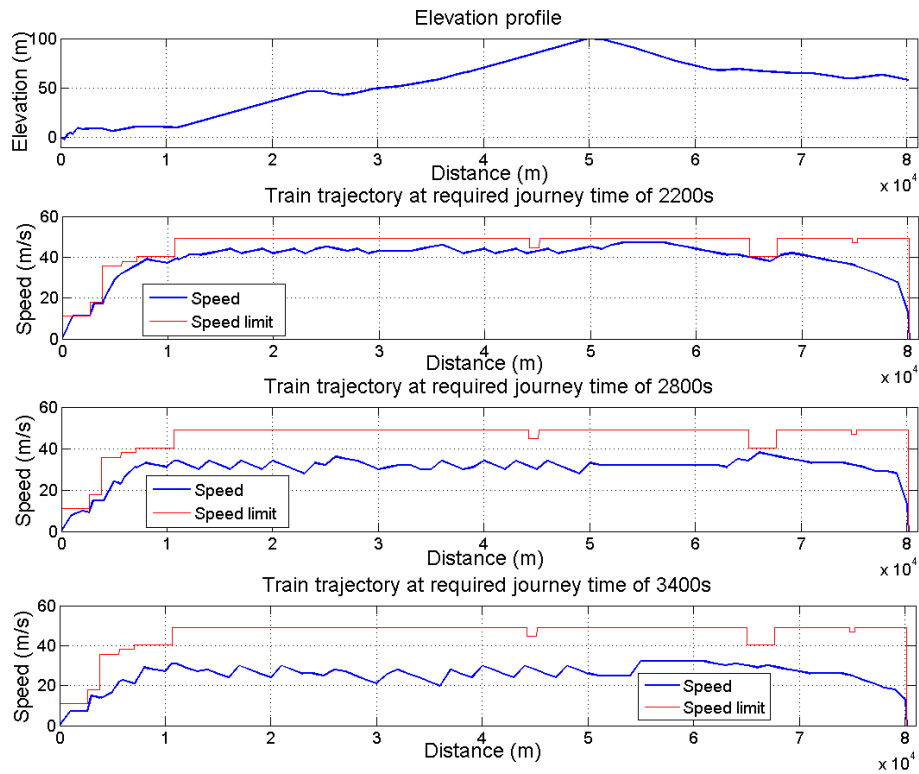


Figure 5.11: Optimised journey trajectories using DP under different journey time conditions

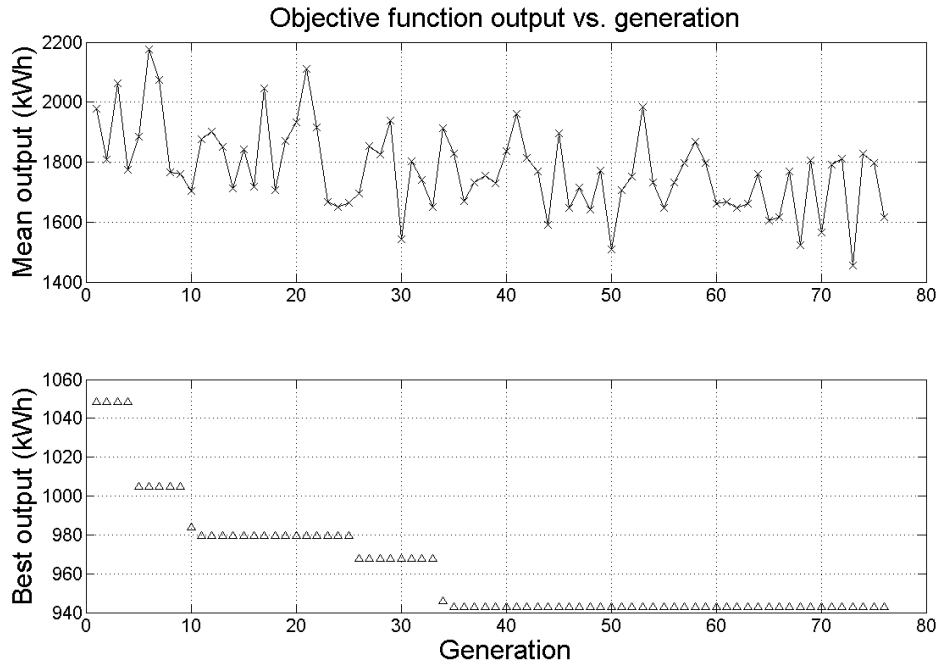


Figure 5.12: The best and mean objective function output for a journey time of 2800 s at each generation for ACO

## 5.7. RESULTS AND DISCUSSION

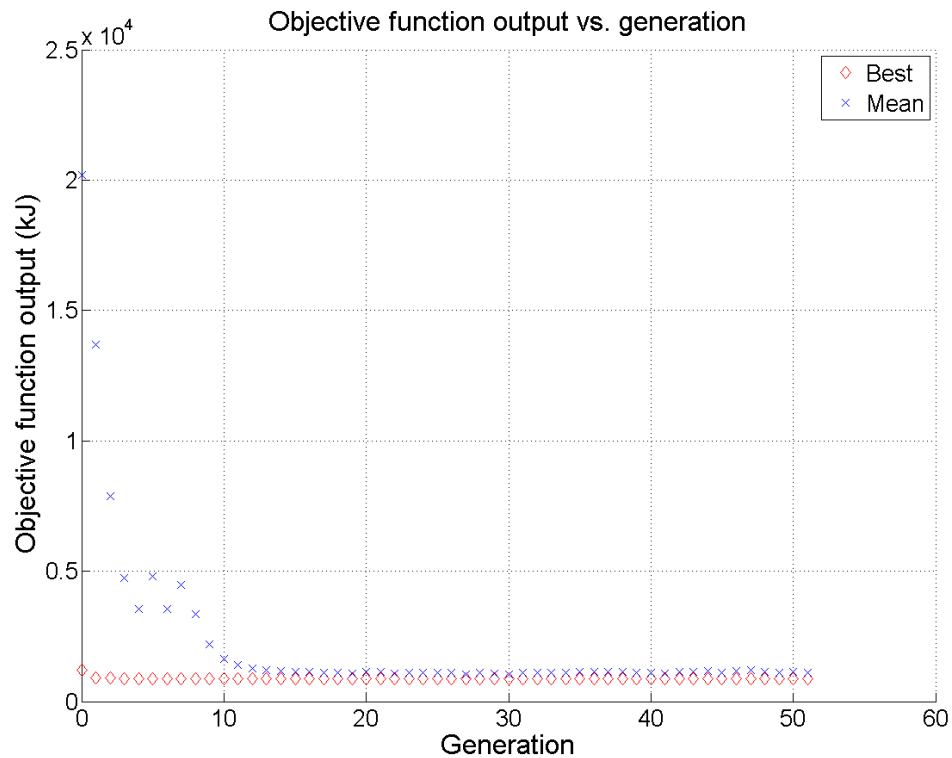


Figure 5.13: The best and mean objective function output for a journey time of 2800 s at each generation for GA

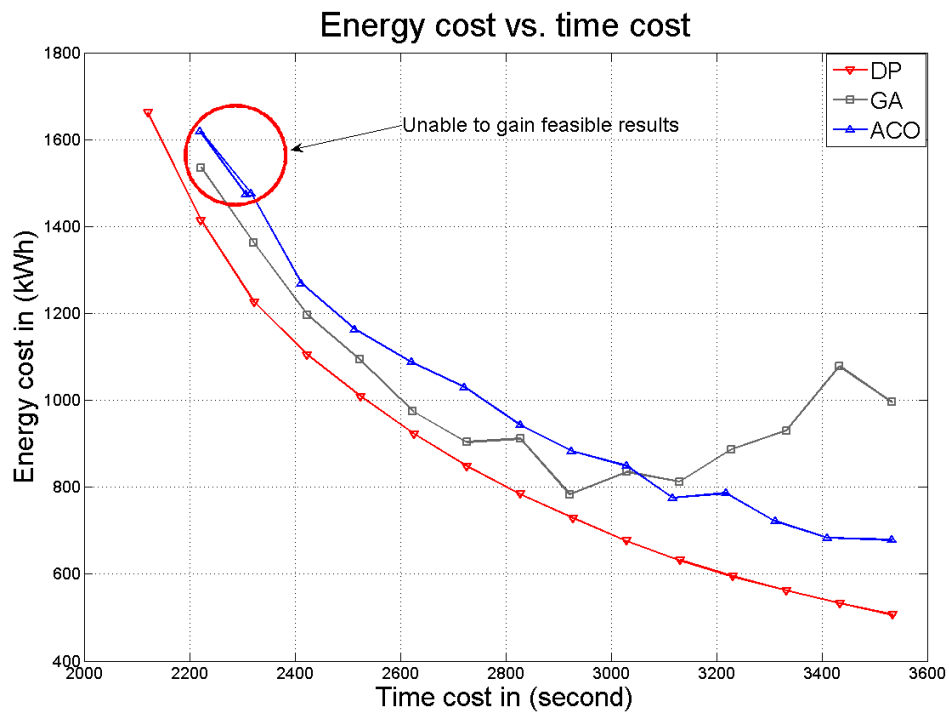


Figure 5.14: Journey energy cost vs. time cost curves using different algorithms

---

## 5.7. RESULTS AND DISCUSSION

---

Table 5.4: Computational time comparison between three algorithms applied on the journey with time requirement of 2800 s for 15 runs

| Algorithms | Mean value | Deviation (%) | Average computational time(s) |
|------------|------------|---------------|-------------------------------|
| ACO        | 946.6      | 16.6          | 552.3                         |
| GA         | 885        | 51.6          | 1550.5                        |
| DP         | 784.6      | 0             | 2431.8                        |

the increase in required journey time. However, strong heuristic information in the ACO prevents the algorithm from obtaining the better solutions with a lower energy cost.

Dynamic programming is able to obtain the best solution amongst the three algorithms, however, the search time efficiency is significantly sacrificed. Since any combination of currently used journey time, currently used energy and currently distance implies a unique state in the searching space, the computational complexity becomes enormous. The DP algorithm demonstrates its robust search capability even for a short journey time, for example 2100 s, a requirement for which the ACO and the GA are unable to find a suitable trajectory as shown in 5.14. When the time requirement is even shorter, an even more rigorous selection of candidate speed at different distance will be required. However, algorithms based on the random Monte Carlo style selection are less capable of reaching the target speed in a relatively short number of iteration steps.

In Table 5.4, a comparison on the computational time between three algorithms are presented. It is demonstrated that more heuristic information will gain results with less deviation quicker. For DP, since the search procedure is determinant, the deviation is zero. The computer used is with Intel dual core CPU 6420, 2.13 GHz and 1.98 GB of RAM.

## 5.8 Summary

In this chapter, single train trajectory optimisation is discussed. Due to the importance of research in this area, several models and algorithms are reported in other research. The method and model proposed in this chapter are inspired by the theory of optimal control, based on which it is concluded that the optimum train trajectory can be obtained from a certain sequence of five train operations. How to choose the sequence is a problem which remains unclear. By using a reasonable approximation of a train speed trajectory over a relatively short distance, searching for the correct sequence is the procedure to determine the speed at different pre-determined positions. A sparse storage model is proposed. Two heuristic algorithms, ACO and GA, are applied based on this model. DP is also used to search for the target speed and the simulation results are demonstrated and discussed. This chapter demonstrates a different way to tackle the train trajectory optimisation problem. Searching for the optimum trajectory should not be only limited to certain train operations, for example, only the train coasting points. Without losing the generality, heuristic algorithms are adopted to avoid the disadvantage of numerical algorithms for optimal control. As the model splits the optimisation into different sub-optimum problems, DP is also applied. Future work will research an increase in the precision of approximation of train trajectories between two pre-determined points and the possibility of changing the searching procedure into an online application.



## **Chapter 6**

# **Optimising the coordinated train operation in a DC railway electric network**

### **6.1 Introduction**

Within an electrical railway network, the manners in which multiple trains are operated affect the energy efficiency. If several trains which are close together demand power at the same time, the amount of current drawn out of a nearby substation will become large. This is usually referred to as the power peak [34–37]. The power peak will give rise to higher total energy loss due to the internal impedance of the substation and the increased  $I^2R$  losses. Furthermore, the power peak will require a larger electrical capacity for the substations even if the average power demand from each train is relatively small.

Different methods are proposed in the literature to reduce the power peak, therefore improving the energy efficiency in a railway electrical network. The train service frequency, nominal braking, acceleration rate, braking effort and other power system configurations are regarded as key factors for power reduction, through which the total power peak could be reduced [39]. An intelligent traction control

method is proposed in [35]. In this study, trains reduce their acceleration power to lower the current peaks in wayside substations. Such a method is reported to achieve a 31% power peak reduction, while not affecting the service quality. More generally, coordinated train control of multiple trains has also been considered and a reduction of the power peak was listed one of the main targets [37].

The synchronism between the operations of trains in an electrical network will affect the network receptivity which is to describe how well the regenerative braking energy can be used. For trains with the capability of regeneration, as reported in [38], synchronism management for the operation of adjacent trains could be considered as a feasible solution to reduce the power peak. A procedure of desynchronising the departure time of trains within a metro system is presented in [36]. Departure time in a the time table is considered as one of the factors which affects synchronism between operations of trains. Reduction of the power peaks will be achieved through better braking power recovery which can be varied by desynchronising the departure time of trains. Similarly, the synchronism can also be changed through the journey time allocations between the inter-station sections. However, such a method will directly involve other issues such as service quality or customer experience. A study presented in [124] has demonstrated that by dynamically varying the journey running time and dwell time at the station, the total service quality and total energy consumption will also be changed. Though no electrical network analysis was involved in this study, the varying of journey time allocations provided a feasible solution to reduce the power peak and total power cost. It has been demonstrated that reduction of the power peak and energy consumption can be achieved by searching for an optimal distribution of the running time of trains along the journey [34]. An optimal distribution of inter-station journey time in an electrical railway network will enhance regenerative braking power recovery due to improved network receptivity, lower the power loss inside the substation due to a reduction of the power peak [26].

To investigate further the relationship between the journey time allocation and

the electrical network energy efficiency, this chapter describes the development of a multiple-train model with regenerative braking in a DC electrical network which is mainly applied regional or suburban railway networks as introduced in Section 2.2.1. With higher density of trains, a DC network is suitable for a study on the co-ordination between trains. Different kinds of DC railway electrical network modelling have been discussed in several articles [139–142]. A multiple-train model in an electrical network is illustrated in both [119] and [143]. The railway DC network model proposed in this chapter achieves the instantaneous states of an electrical network by the LF method in association with the Least Square Minimisation (LSM) method. The study proposed searches for the optimal journey time allocation for each inter-station run to maximise multiple objectives including the energy efficiency and the service quality. The service quality is measured by the time consumption while energy efficiency will be affected by the energy loss within the substations and the electrical circuits, both of which are related to the power peak. A GA is applied to search for the optimal allocations and results achieved from this method have been bench-marked by the one achieved through a Brute Force (BF) search.

## 6.2 DC railway network modelling

### 6.2.1 DC electrical railway network

In a DC electrical railway network, all the feeding routes are regarded as resistive [140]. A simplified network is shown in Figure 6.1.

A similar DC traction load flow study including earthing models has been covered in [142]. For the sake of simplicity, the leakage resistors have been ignored in this study. A diagram shown in Figure 6.1 demonstrates a typical DC electrical network with a running train [144].

In Figure 6.1,  $R_{sub}$  is the substation internal resistor. It is regarded as constant

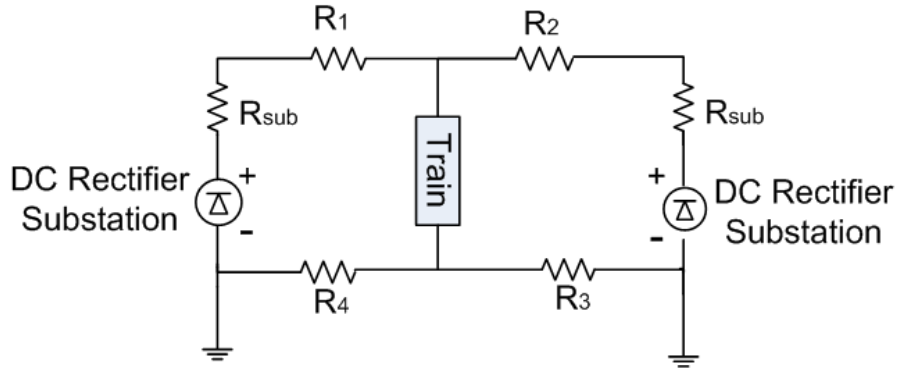


Figure 6.1: A simplified DC railway electrical network diagram

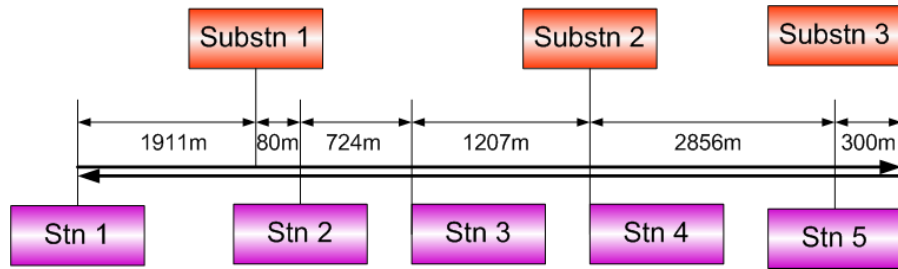


Figure 6.2: Physical location arrangement of stations and substations in a DC railway network

and “0.05  $\Omega$ ” is used in this study.  $R_1$ ,  $R_2$ ,  $R_3$  and  $R_4$  are the feeding line resistors which are dependent on the location of train. The resistant rate adopted in this study is 40.61  $\mu\Omega/\text{km}$ . All the value used is typical for a DC railway [22].

The DC rectifier substation does not allow the current to go back into the power supply system and it is regarded as a constant voltage source. The output voltage value in this chapter is adopted as 780 V DC.

The train will be regarded as a power source during braking and a power consumer in the normal motoring mode. Output power is positive when the train is regenerating and negative while the train is absorbing the traction power from the network.

Figure 6.2 shows the physical location arrangement of stations and substations in a typical DC network. Note that the route consists of two parallel tracks.

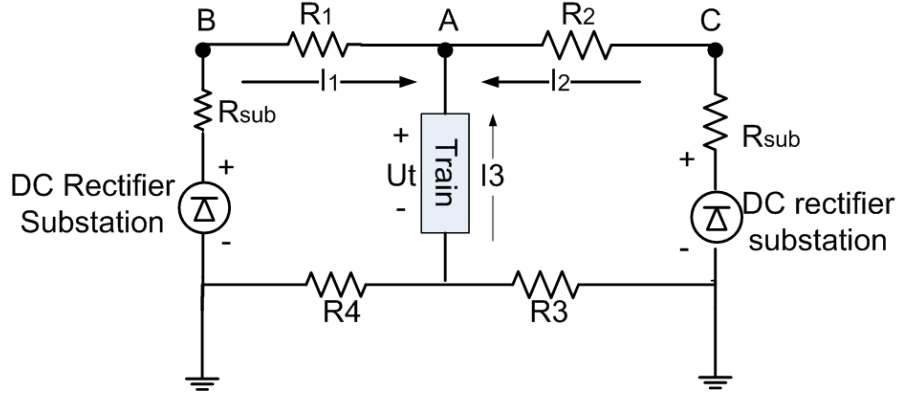


Figure 6.3: Nodal analysis

### 6.2.2 DC electrical analysis

Electrical network analysis is a procedure to produce a set of equations which can describe the internal electrical relationship between all the junctions in the network. The procedure is based on fundamental electrical law, i.e. Kirchhoff's circuit laws. Nodal analysis and mesh analysis are two ways to analyse a DC electrical network. In this study nodal analysis are used to analyse the network [145]. Nodal Analysis (NA) is based on the principle of conservation of electrical charge. *The principle of conservation of electric charge implies that: at any node or junction in an electrical circuit, the sum of currents flowing into that node is equal to the sum of currents flowing out of that node.*

In Figure 6.3, using nodal analysis, it is concluded that all the current which goes into junction A must equal zero, as shown in Eqn. 6.1.

$$I_1 + I_2 + I_3 = 0 \quad (6.1)$$

Consequently:

$$\frac{U_B - U_A}{R_1} + \frac{U_C - U_A}{R_2} + I_3 = 0 \quad (6.2)$$

Because  $I_3$  is the terminal current of the train and  $U_A$  is the terminal voltage of

the train:

$$I_3 = I_t \quad (6.3)$$

$$U_A = U_t \quad (6.4)$$

Rewrite Eqn. 6.2:

$$\left(\frac{1}{R_1} + \frac{1}{R_2}\right)U_t + \left(-\frac{1}{R_1}\right)U_B + \left(-\frac{1}{R_2}\right)U_C = I_t \quad (6.5)$$

In Eqn. 6.5, the term “ $\frac{1}{R_1} + \frac{1}{R_2}$ ” is referred to as self admittance of junction A, and “ $-\frac{1}{R_1}$ ” and “ $-\frac{1}{R_2}$ ” is the mutual admittance between A and B, B and C respectively [145]. The current  $I_t$  is also called the injection current for junction A when  $I_t$  is positive in direction, as shown in Figure 6.3.

### 6.2.3 Single train simulation using coasting control

Single train simulation in this study is done by time based modelling, as discussed in Section 4.3.1. A typical urban railway vehicle with a maximum power output of 656 kW has been modelled. The specific tractive effort and the resistive curve is shown in Figure 6.4.

An Ordinary Derivative Equation (ODE) solver is applied to obtain various vehicle states along an inter-station journey. Meanwhile, in order coasting control is adopted to meet the dynamic time allocation for each inter-station journey. Figure 6.5 illustrates an idea of vehicle state switch behind the train running trajectory optimisation using coasting control.

In Figure 6.5, a vehicle state consists of current vehicle speed, distance, time and energy consumed since the initial state. Vehicle states have been created using both motoring operation and coasting operation. The energy consumption of states with duplicate speed, distance and time will be compared first. States with more energy consumption will be eliminated. States with unrealistic elements, such as

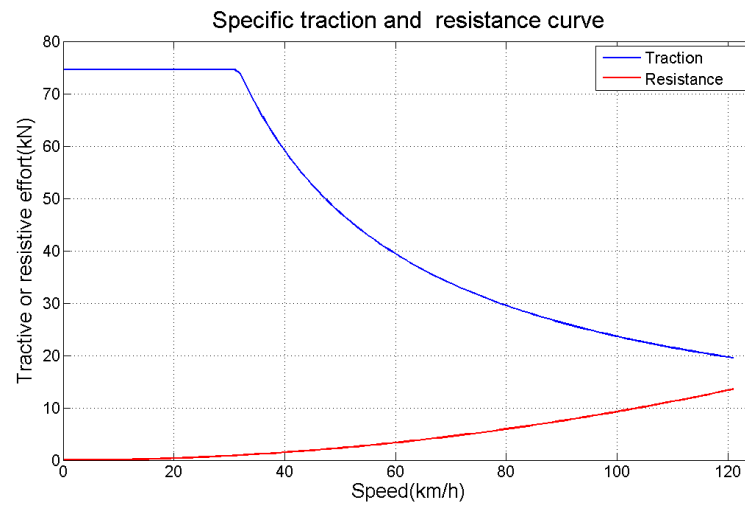


Figure 6.4: Maximum tractive effort and resistance curve of a typical urban railway vehicle.

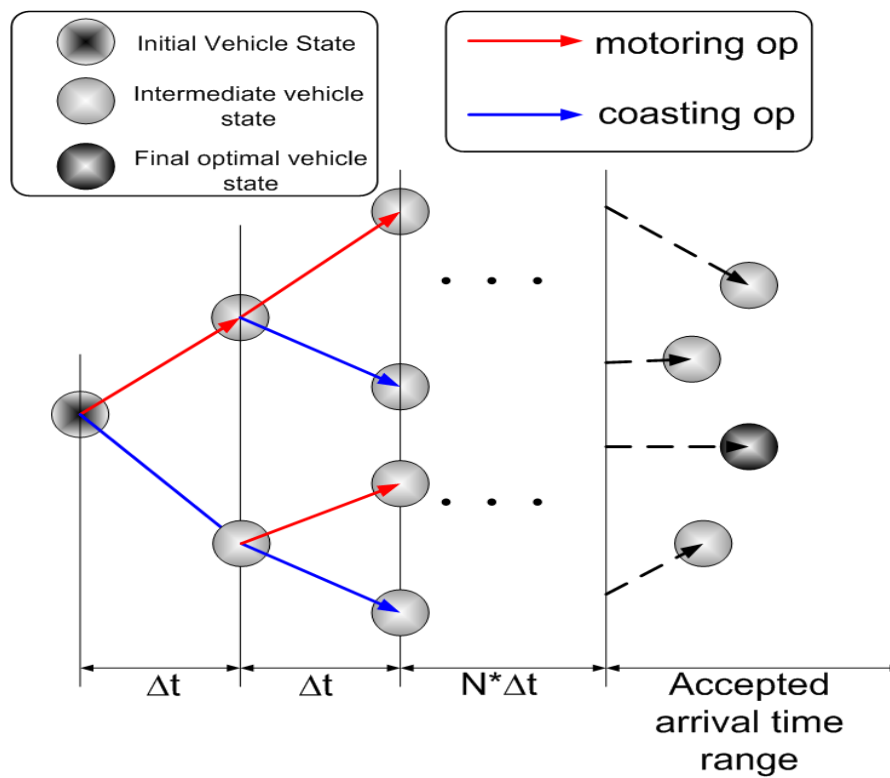


Figure 6.5: Vehicle state switch using coasting control.

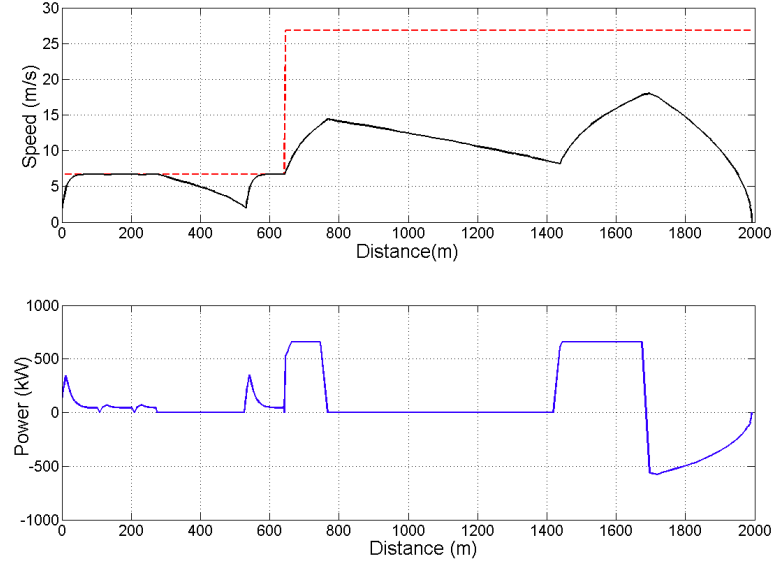


Figure 6.6: Optimised train running trajectory and instant power output using coasting control strategies for journey between station 1 and station 2

too much time consumption, undesirably low speed etc will also be ruled out. A similar searching procedure, also covered in Section 5.6 is applied to search for the optimal coasting point to meet the dynamic journey time allocation. Figure 6.6 shows a searched train running trajectory and instantaneous power output using coasting control on the 1991 m long journey from station 1 to station 2.

Single train simulation produces a running profile with vehicle states and corresponding instant power demands. The produced running profile should meet the time constraint imposed by the dynamic journey time allocation. The instant power output will be used later to simulate the total DC network, from which the energy efficiency of the total network can be determined.

#### 6.2.4 Iterative power flow calculation

To achieve this, an additional equation is added to account for the instantaneous train power:

$$P_t = U_t I_t \quad (6.6)$$



where  $P_t$  is the power demand for each train.

The junctions outside the substation and junctions where the trains are located will also be accounted for. There are 3 substations, and several trains along the forward and backward routes. The number of trains  $N_t$  on each route is determined by the service headway  $T_h$  (which is the departure time difference between two subsequent trains) and total journey time which is the sum of all inter-station journey time allocations  $T_{total}$ .

$$N_t = \left\lfloor \frac{T_{total}}{T_h} \right\rfloor \quad (6.7)$$

where  $\lfloor X \rfloor$  is the greatest integer less than or equal to “X”. It is noted that  $T_{total}$  is different for the routes of different directions.

$$\begin{bmatrix} A_1^s & \cdots & A_{1(M+N+3)}^m \\ \vdots & \ddots & \vdots \\ A_{(M+N+3)_1}^m & \cdots & A_{M+N+3}^s \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_{M+N+3} \end{bmatrix} = \begin{bmatrix} I_1 \\ \vdots \\ I_{M+N+3} \end{bmatrix} \quad (6.8)$$

Assume that there are N trains moving forwards, i.e. from Station 1 to Station 5 for a period of headway time and M trains travelling backwards, i.e. from station 5 to station 1. In Eqn. 6.8, subscripts 1 to 3 denote the junctions outside of substations. Subscripts 4 to N+3 denote nodes of the forward trains, and subscripts N+4 to M+N+3 denote the backward trains. The superscript “s” implies self-admittance and “m” implies mutual admittance between two nodes. For mutual admittance, the subscripts denote the trains.

An admittance matrix is required to perform the NA. This is done by calculating the resistance between two junctions using the resistive rate. If two junctions are not adjacent to each other, the resistance in between is set as infinity and as a result, the mutual admittance in between will be zero. It is obvious that the admittance is a sparse matrix.

Eqn. 6.8 is a non-linear equation as the relationship between  $U_i$  and  $I_i$  is non-

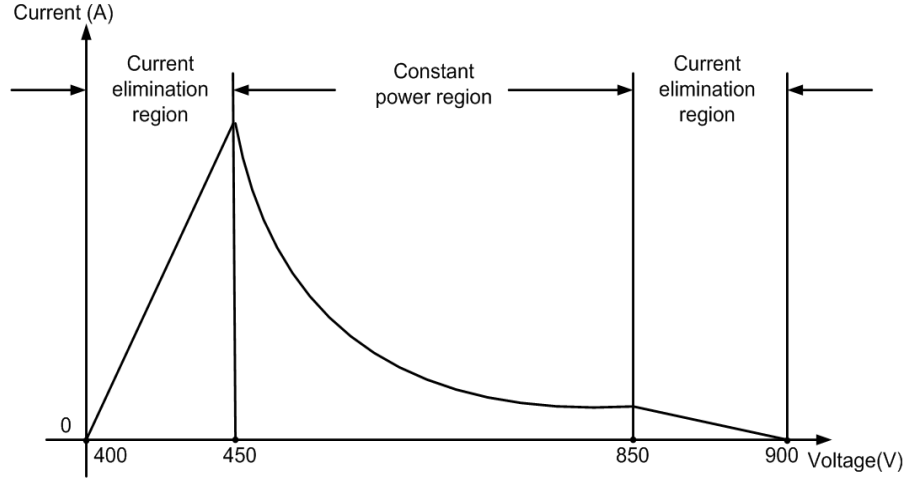


Figure 6.7: Non linear relationship between the terminal voltage and current of a train

linear. Expected problems may arise due to the surplus voltage during regeneration. However, in practice, when surplus voltage occurs, a bank of resistors will be connected within the train power supply system to consume the surplus power and clamp the output voltage at a certain level. Meanwhile, for substations, whenever the voltage outside of a substation is over 780V, the current out of the substation is virtually zero to reflect the non-reverse conductivity of diodes. Figure 6.7 and Figure 6.8 demonstrate this non linear relationship between the terminal voltage and current for trains and substations. It is noted that the current will be kept as constant as long as there is an injected current from the substation. No current will be injected if the terminal voltage is over 780V and the substation is virtually shut down at this moment.

In Figure 6.7, 3 regions based on the level of voltage are defined. Between the range [450,850] the relationship between voltage and current will be determined by Eqn. 6.6. However, for voltages over 850 V and below 450 V, the current will be linearly reduced to zero.

In Figure 6.8, the constant injected current can be calculated using Eqn. 6.9.

$$I_{sub} = \frac{U_{sub}}{R_{sub}} \quad (6.9)$$

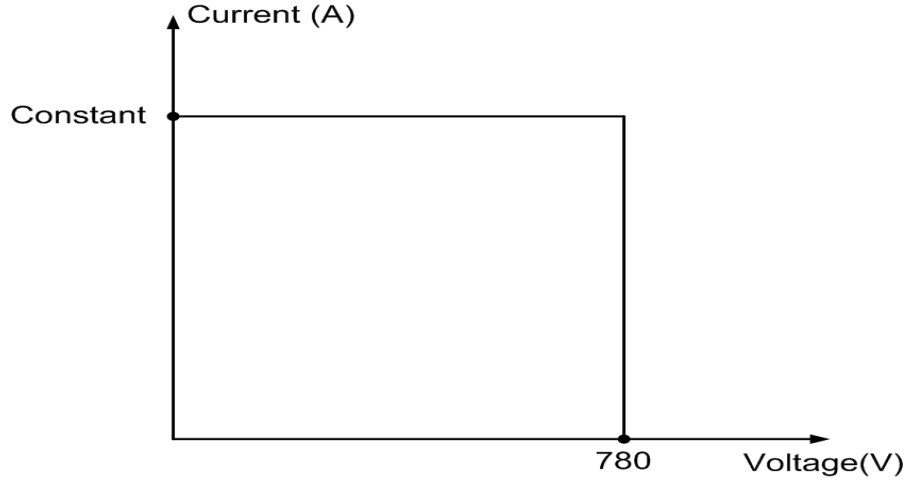


Figure 6.8: Non linear relationship between the terminal voltage and injected current out of substation.

where,  $U_{sub}$  is a constant value 780V and  $R_{sub}$  is  $0.05 \Omega$  [22].

To obtain the solutions in terms of instant current and voltage, the LSM method is used in this study. The set of scalar  $U_i$  are dependent on scalar  $I_i$ .  $F_a$  is used to denote the electrical relationship, defined in Eqn. 6.8.

$$I_a = F_a(U) \quad (6.10)$$

The subscript “a” implies that the corresponding matrix is an admittance matrix. Meanwhile, as shown in both Figure 6.7 and Figure 6.8, the non-linear relationship between  $U$  and  $I$  can be presented in Eqn. 6.11.

$$I_n = F_n(U) \quad (6.11)$$

A distinction between  $I_a$  and  $I_n$  may exist if  $U$  is not a solution of Eqn. 6.11.  $E^i$  denotes the absolute difference between “ $i_{th}$ ” elements of  $I_a^i$  and  $I_n^i$ .

$$E^i = |I_a^i - I_n^i| \quad (6.12)$$

Based on Eqn. 6.12,  $E$  is used to denote the total square sum of each  $E^i$ :

$$E = \sum_{i=1}^{M+N+3} (E^i)^2 = \sum_{i=1}^{M+N+3} (I_a^i - I_n^i)^2 = \sum_{i=1}^{M+N+3} (F_a(U^i) - F_n(U^i))^2 \quad (6.13)$$

If, by any means, the minimum value of  $E$  can be found, the set of scalar  $U$  will be sufficiently close to the real solution. LSM is applied to find the set of  $x$  to minimise Eqn. 6.13.

## 6.3 Methodology

### 6.3.1 Objective function

As introduced in Section 6.1 the system cost function is based on two considerations.

Firstly, the service quality of the journey time should be taken as the optimisation target. It is assumed that people prefer a shorter journey time. The shorter the journey time is, the better the service quality will be and higher fitness value will be obtained.

For any inter-station journey, the service quality is calculated using the equations as follows:

$$Q_s = W_i M_i \quad (6.14)$$

$$W_i = T_i / T_{total} \quad (6.15)$$

$$M_i = \begin{cases} 1 & \text{mode} = \text{fast} \\ 0.5 & \text{mode} = \text{normal} \\ 0 & \text{mode} = \text{slow} \end{cases} \quad (6.16)$$

where:

- $Q_s$  is the total quality of service;
- $W_i$  is the weight of the current journey;
- $M_i$  is the fitness value for the journey time.

It is noted that inter-station journey running is categorised into three running modes: fast-mode, normal-mode and slow-mode. In order to provide equivalent service difference between each mode, 1, 0.5 and 0 are used in Eqn. 6.16. Each mode corresponds to a type of journey with a certain amount of journey time cost. Intuitively, the journey in fast-mode consumes less time. Meanwhile, the ratio that the inter-station journey time takes in the total journey time determines the importance it takes in the entire journey, therefore it is also imposed in the journey quality evaluation. In other words, the larger portion it takes in the total journey time, the more important an inter-station will be and the more easily it will affect the total journey quality.

The energy efficiency is then taken into account. The evaluation of the energy efficiency can be achieved by the following equation:

$$E_s = \frac{E_{out}}{E_{in}} \quad (6.17)$$

where:

- $E_{out}$  : the total output energy, i.e. the total terminal energy consumption for all trains;
- $E_{in}$  : the total input energy from the substations, i.e. the total injected energy from substations.

It is noted that if more power is consumed in the internal impedance of a substation, lower energy efficiency will be achieved.

The weightings for both service quality and energy efficiency which contribute

to the total evaluation function is shown as follows.

$$\varepsilon = W_1 Q_s + W_2 E_s \quad (6.18)$$

where,  $W_1, W_2 \in [0, 1]$ ,  $W_1 + W_2 = 1$  and  $\varepsilon$  is the total evaluation function output. In this study  $W_1 = 0.4$  and  $W_2 = 0.6$  are arbitrary adopted to reflect individual importance of two elements.

### 6.3.2 Journey time allocation

It is understood that journey time allocation will affect both service quality and network energy efficiency. The procedure of producing the dynamic journey time allocation is introduced as follows.

Firstly, all inter-station operations are simulated using a flat-out driving style. Based on the results of the simulation, the minimum time  $T_{min}$  will be recorded for each simulated inter-station journeys.  $T_{min}^1, T_{min}^2, T_{min}^3$  and  $T_{min}^4$  denote the minimum time for the four inter-station journeys for the forward route where a train runs from station 1 to station 5 and  $T_{min}^5, T_{min}^6, T_{min}^7$  and  $T_{min}^8$  to denote the minimum time for 4 inter-station journeys on the backward route where a train runs from station 5 to station 1. The minimum time costs for all the inter-station runs are represented as  $[T_{min}^1, T_{min}^2, \dots, T_{min}^5, \dots, T_{min}^8]$ . Based on the minimum journey time, for each inter-station the journey time at each running mode can be calculated using equations shown in Eqn. 6.19. The journey times selected for each mode of journey are typical values for regional railway services.

$$T_{alloc}^i = \begin{cases} 1.1T_{min}^i & mode = fast \\ 1.2T_{min}^i & mode = normal \\ 1.3T_{min}^i & mode = slow \end{cases} \quad (6.19)$$

Furthermore, in order to represent various journey time allocations, each run-

ning mode for each inter-station journey is coded using a characteristic number. In this study, number “1” is used to represent a fast mode operation for an inter-station journey, number “2” denotes a normal mode operation and number “3” denotes a slow mode operation. Therefore, an 8-member array consisting of 8 characteristic numbers will represent a journey time allocation since the index of numbers in the array suggests the journey it is corresponding to while each characteristic number provides the information of the journey time. For example, array [1, 1, 3, 3, 1, 2, 3, 2] will represent a time allocation for 8 inter-station journeys shown as follows:

$$[1.1T_{min}^1, 1.1T_{min}^2, 1.3T_{min}^3, 1.3T_{min}^4, 1.1T_{min}^5, 1.2T_{min}^6, 1.3T_{min}^7, 1.2T_{min}^8]$$

It is noted that a combination of journey time selections suggests a set of journey profiles which is pre-obtained from a single train simulation using coasting control. Such a pre-obtained journey profile is then used to calculate the DC electrical status within a headway time. Using Eqn. 6.14 and Eqn. 6.17, the evaluated value for such journey time allocation can be calculated.

### 6.3.3 The genetic algorithm

A GA is used to search for the optimum journey time allocation for each inter-station journey. The results are compared with the one obtained by the brute force searching method. Each solution for the evaluation function is encoded as a chromosome in a population. The process of the GA is illustrated in a diagram shown in Figure 6.9.

In Figure 6.9, the solution of the evaluation function is encoded as a chromosome. The evaluation function is comprised of both a service evaluation and an energy efficiency evaluation. Because the evaluation is to be maximised, it can be used as the fitness function directly. The fitness function calculates the fitness value for each solution in one generation based on the ranking in all the solutions.

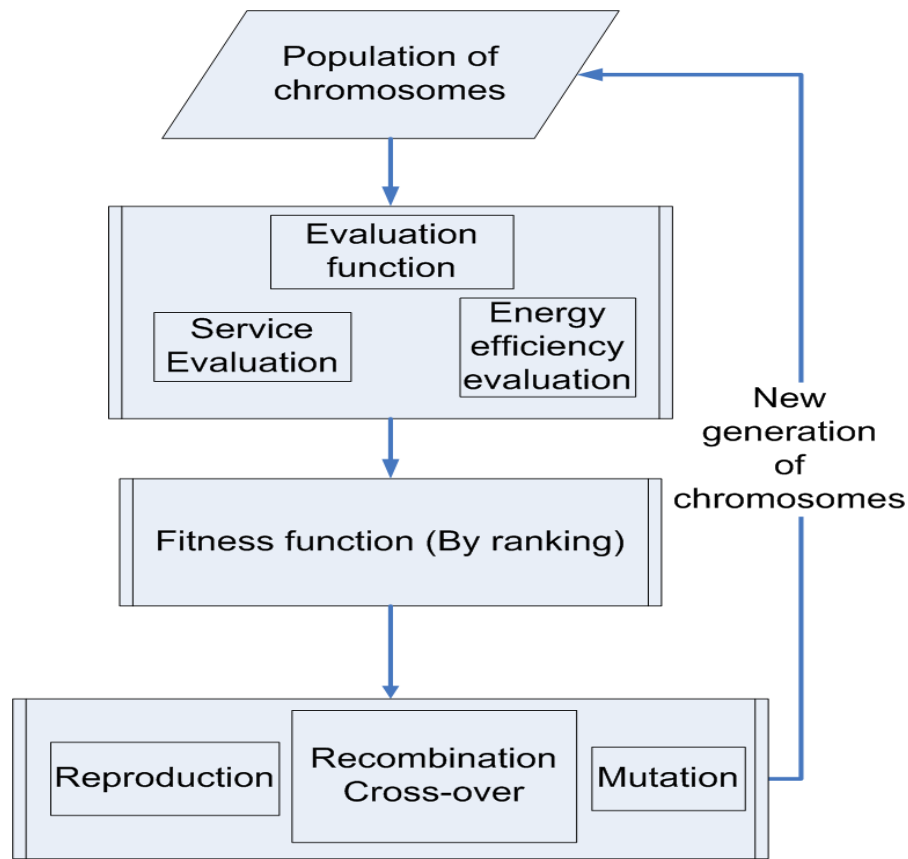


Figure 6.9: Genetic algorithm implementation chart flow

The fitness function makes sure that the better evaluated solutions are selected with a higher possibility to be operated for next generation. After the GA operations, including reproduction, recombination and mutation, a new generation of chromosomes is produced and the whole process is taken iteratively until the termination criterion is met as introduced in Section 3.4.2.

Reproduction is for the best solution(s) to have a chance of going directly into the next generation. In this study, only one best solution is selected for the next generation. The cross-over recombination process was taken between two chosen chromosomes. The possibility of two chromosomes being selected is based on the ranking of respective evaluated values amongst all solutions. Intuitively, the best solution will have the highest possibility of being selected for recombination. After different test search, the mutation rate is set initially at 0.2, which means that for 20 chromosomes, around 4 chromosomes are produced from each mutation for better



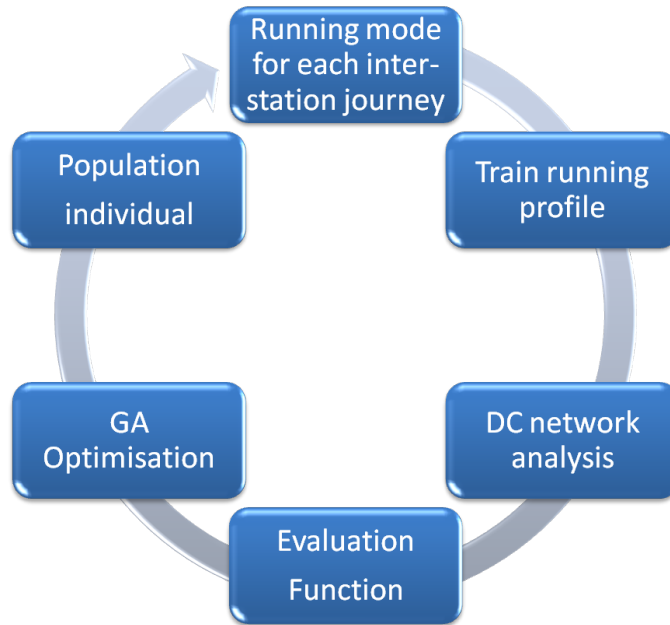


Figure 6.10: Entire optimisation procedure using Genetic Algorithm.

search performance,. However, it will help to make a broader exploitation within the solution space. Mutation tends to happen more frequently for better solutions and only one number is allowed for change.

More generally, the entire optimisation process, including all the afore-mentioned modelling and calculation can be illustrated using Figure 6.10. In each iteration, an entire journey profile will be generated using the genotype information which is also a feasible solution. A DC network analysis is then done to evaluate the fitness function using the journey profile and results of the evaluation will guide the GA continue to search for better solution.

## 6.4 Results and discussion

All journey running curves for different journey times are stored in a look-up table and the running trajectories are assumed to be stationary for the simplicity of this study. One journey time allocation can be determined by a set of characteristic numbers and each number will correspond to one typical train running trajectory in the look-up table. As a result, the entire train running trajectory for the whole

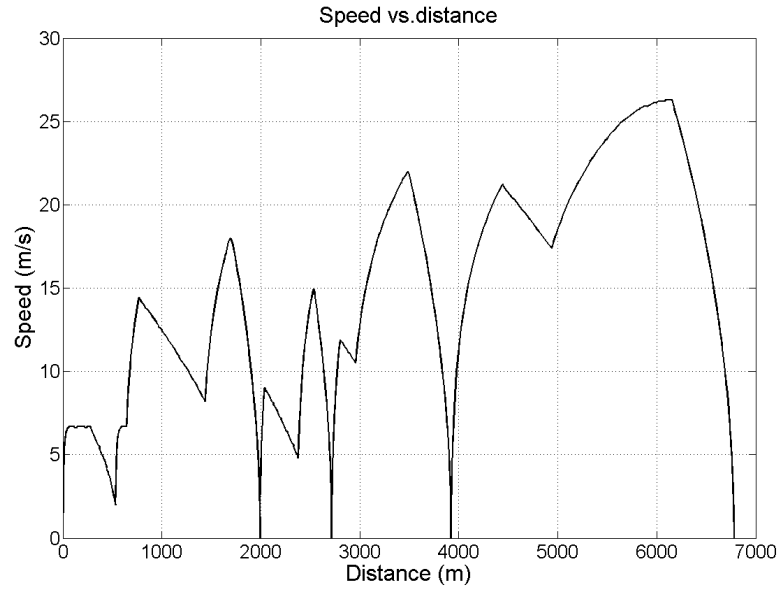


Figure 6.11: Train speed vs. time

journey can be obtained and will be used to calculate the total energy efficiency and service quality. Figure 6.11 demonstrates a typical train running trajectory vs. distance from station 1 to station 5.

It is assumed that trains are operated automatically and each train runs exactly along the running curve under the hypothetical condition that strict constant headway time and constant station dwell time (30 s) are maintained. By regarding the train operation as a periodic process, each running train will be distributed along the routes with constant headway time. Meanwhile, the instantaneous position of each train can be determined using the relationship between journey time and distance shown in Figure 6.12. Five stations have been pin-pointed with no change in distance in a 30-second dwell time.

Subsequently, the DC electrical network characteristic will be determined at each position along an inter-station journey, because not only the speed but also the instant power output of trains can be determined, as shown in Figure 6.6. The instant status of the DC electrical network is calculated using DC electrical analysis. Figure 6.13 and Figure 6.14 present the instant voltage and current output from the three substations. Note that 240 s is selected as a typical headway time for regional

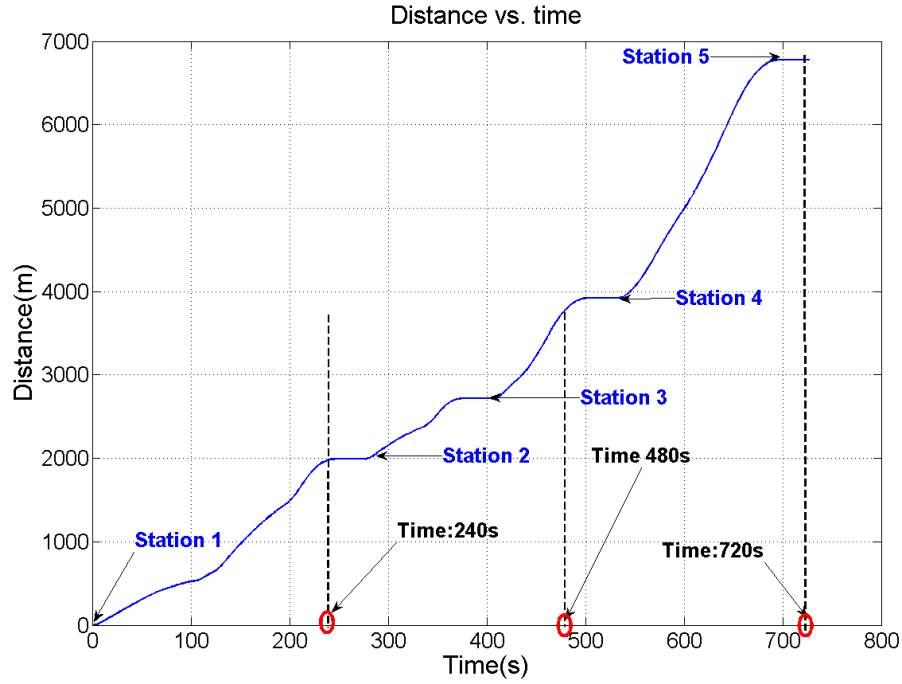


Figure 6.12: Train distance vs. time

railway network. All trains follow the same trajectory, with a headway time of 240 seconds and initially there are trains at 0 s, 240 s, 480 s, and 720 s respectively. This also is highlighted using red circles in Figure 6.12.

The introduction of fluctuation to the journey time will not only directly affect a single train power performance, but also the trains nearby within the electrical network. The energy efficiency of the total network will subsequently be affected. Because the change in journey time will significantly change the synchronous time for train braking operations in an inter-station operation, the total energy recuperation rate will vary accordingly.

Figure 6.15 shows one of the GA evolutions. It is noted that the best solution of [2, 1, 1, 1, 1, 1, 1, 1] and 0.85 evaluation function output has been achieved with the energy efficiency of 0.82. This implies that in order to achieve the higher efficiency and better service quality at the same time, it is good to increase the journey time slightly from fast mode operation to a medium fast mode for the first inter-station journey. Fast mode is preferable for train operations on other inter-station journeys.

## 6.4. RESULTS AND DISCUSSION

---

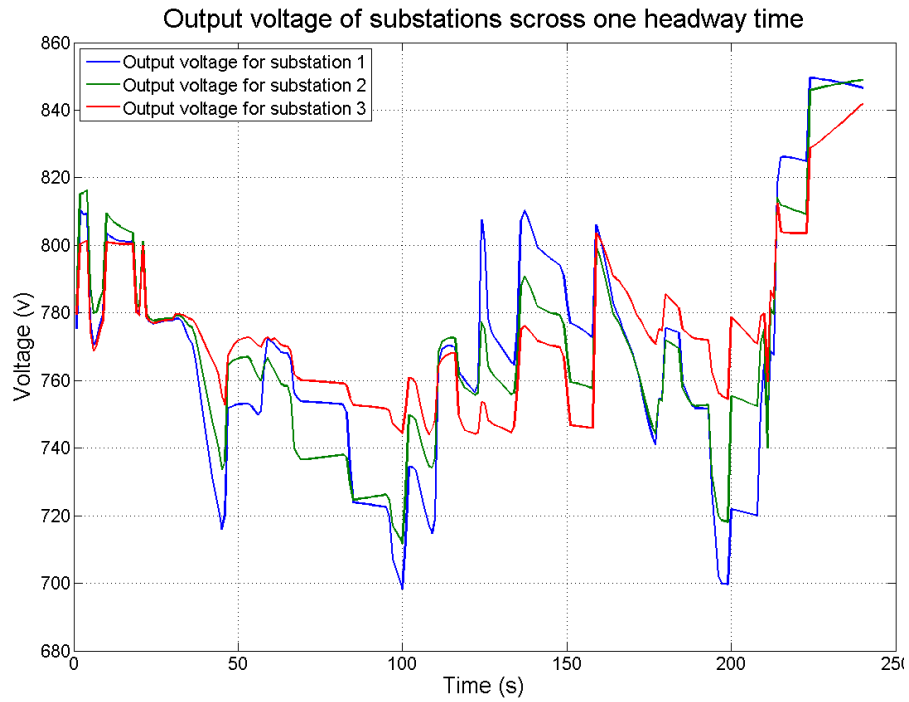


Figure 6.13: Instant voltage output across a headway time 240s.

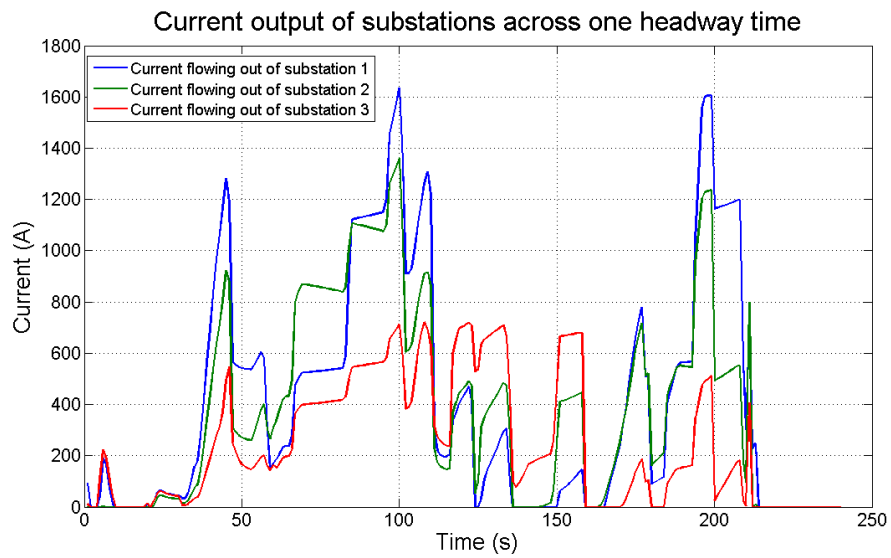


Figure 6.14: Instant current output across a headway time 240s.

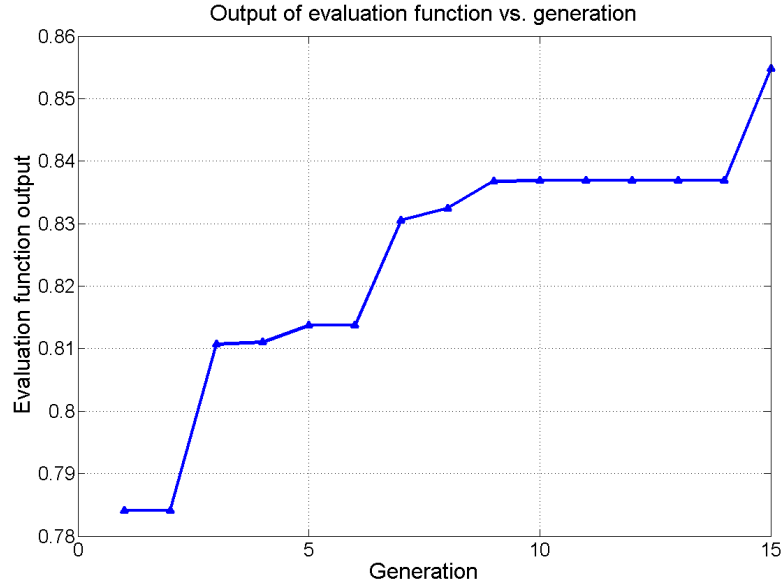


Figure 6.15: Evolutionary fitness function output vs. generation in the GA

The optimal result obtained by the GA is compared with the one obtained from the BF method. By searching all of the possible combinations of running mode, which is  $3^8 = 6561$ , it is finally found that the best time allocation mode vector is as follows [1,1,1,2,1,1,1,1]. In this global optimal solution, only the fourth inter-station run is allowed to increase its journey time to normal running. Figure 6.16 demonstrates the distance vs. time for four journeys with different inter-station journey time allocations. Along with two searched journeys, one by GA in black and one by BF in red, journeys with the longest and shortest journey time allocation are also included.

## 6.5 Summary

In this chapter, the journey time allocation within a DC railway network is discussed. The simulation on a multiple-train system within a DC electrical railway is demonstrated. This study firstly produces a look-up table which contains the optimised single train running trajectories for various journey time requirements. A set of journey time lengths are then selected and the electrical energy efficiency

---

## 6.5. SUMMARY

---

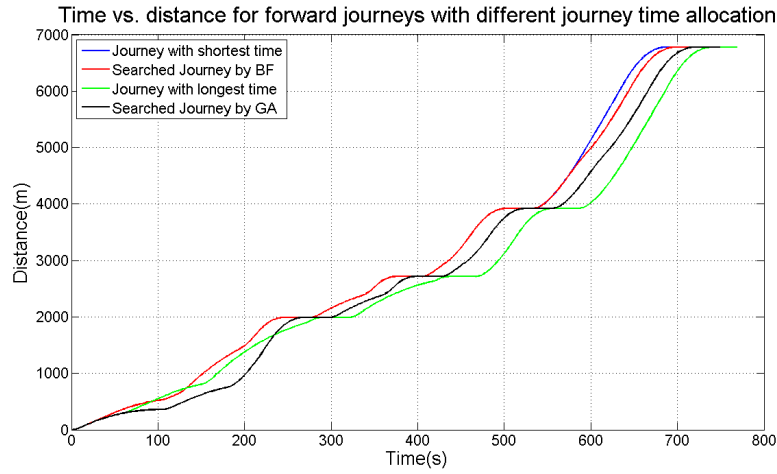


Figure 6.16: Distance vs. Time for forward journeys with different journey time allocation

and service quality are calculated. The evaluation function output can be consequently produced based on the journey time allocation. Finally, GA is used to find a good solution to identify the respective journey time for each inter-station run. The researched results are bench-marked using the global optimal result by a BF method. The results suggest that increasing part of the inter-station journey time reserve will result in the best trade-off between service quality and total energy efficiency of the network.

## **Chapter 7**

# **A power management strategy for a Diesel Multiple Unit train**

In the driving system of a hybrid electric vehicle, regenerated energy can be stored in energy storage devices on board [146] [59] [121]. This application introduces the concept of power management strategy to manage the different power systems on board. In this thesis, taking inspiration from the hybrid electric vehicle, the potential to apply power management strategies to a more commonly deployed type of train, i.e. the DMU train is investigated. The use of multiple power systems in railway applications is common to modern rolling stock. Prime movers are usually distributed along the length of the train, and are able to provide distributed traction. If the traction power for such vehicles comes from individual diesel engines, the supervisory controllers for each multiple unit may be decoupled and operated independently and this can potentially save energy.

The traditional traction system of a DMU train consists of several similar or identical power systems. Multiple unit trains usually operate individual engines in a synchronised manner. This means that engines can often operate well outside their optimum operating region. For example, a set of engines may all work with a low power output resulting in an inefficient operation. The central hypothesis of the current thesis is that by decoupling the engines and operating them individually,

the overall efficiency can be improved using some optimisation techniques.

This chapter will firstly review the current technologies applied in the power management of Hybrid Electric Vehicle (HEV) and then take a typical DMU train as a case study target and demonstrate the application of advanced power management strategies in such a railway vehicle. DP will be used to develop the off-line global optimisation strategy, and the results will then be used to develop the on-line rule-based strategy [48, 49].

### **7.1 Review of power management strategies**

The idea of hybridisation of a propulsion system was originally conceived from the motivation to extend the working range of electric automobiles [40]. However, there are numerous additional benefits of these systems. A hybridised design with both prime mover (Internal Combustion Engine or Fuel Cell) and on-board energy storage device is often able to utilise a significantly down-sized prime-mover. The energy storage device can enable the prime mover to operate within its optimum efficiency range and also capture braking energy [5, 46]. There is a large body of research which has focused on elevating the efficiency of components, such as power electronics devices [147, 148] and batteries [43–45].

In order to take full advantage of hybridisation, effective power management strategies are necessary. Four key goals of a hybrid system are summarised below [47]:

- Maximum fuel economy;
- Minimum emissions;
- Minimum system costs;
- Good driving performance.

The power management strategy needs to consider: the optimal engine operating region, engine dynamics, minimum engine speed, battery state of charge,

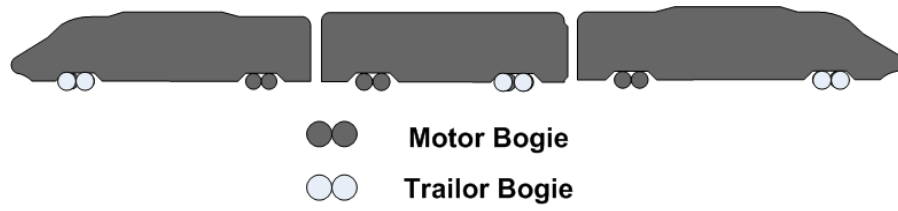


relative power distribution, *etc* [149]. These strategies can be primarily divided into two categories: rule-based and optimisation-based [149].

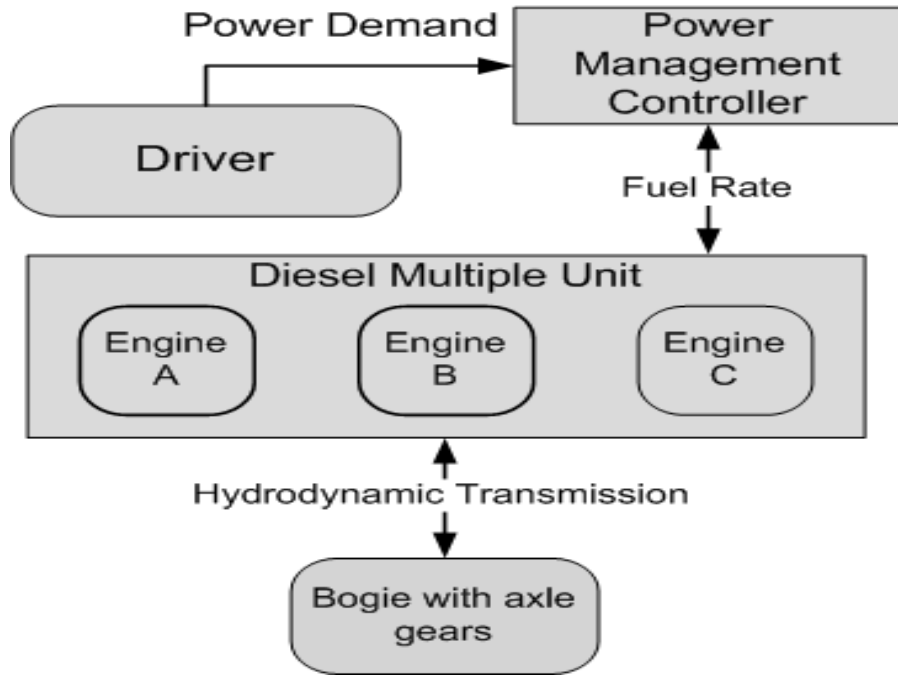
A rule-based strategy consists of sets of if-then rules in an expert system. These sets of if-then rules can be obtained from heuristics, human experience or simulation results. A rule-based strategy attributes its notable advantages to having no requirement for the future journey profile to be known, and it is also suitable for on-line applications. The main idea behind a rule-based strategy for HEVs is “Load leveling”. This method tries to keep the Internal Combustion Engine (ICE) operating in its optimal region. The difference between the output of the Internal Combustion Engine (ICE) and the demands of the driver will be met by the energy storage device [150]. Recent work has shown that Equivalent Consumption Minimisation Strategy (ECMS) can be used to determine an effective rule-based strategy to achieve a fuel saving [151]. Within the scope of rule-based strategies, fuzzy rule-based strategies [152–154] as a robust control method, are suitable for highly non-linear, multi-domain, time-varying systems such as HEV propulsion systems.

Several optimisation-based control strategies for the power management of HEVs have already been proposed. These control strategies could generally be categorised into two groups: global off-line optimisation and on-line optimisation [155, 156].

Global optimisation methods are rarely suitable for on-line implementation. DP [59, 157] is a global optimisation method, but suffers from a long computation time for complex problems. Additionally pre-acquisition of the future data eliminates the possibility of online application. However, where less computational time is not a priority and multiple stage decision making process is achievable, DP still makes a significant contribution. In addition, with its global optimisation characteristic, DP can be applied in the design phase or can be used to improve online power management strategies. The other optimisation methods such as Non-Linear Convex Programming [158], Genetic Algorithms [159, 160], and Optimal Control



(a) Train configuration overview



(b) Train drive system overview

Figure 7.1: A typical 3 coach DMU train

Theory [161,162] have all been applied to develop the power management strategy of HEVs.

## 7.2 Typical DMU train

Figure 7.1(a) shows a schematic of the traction system arrangement of a DMU train. A diesel engine which has the maximum output power of 560 kW is installed in each car. The sizing of the engines used in this work has been closely based on realistic vehicles, for example, the British Rail (BR) Class 185. Figure 7.1(b) is a simplified overview of the driving system [163].

Table 7.1 presents some technical data for this type of vehicle.

Table 7.1: Typical technical data for DMU type vehicles

|                         |                                       |
|-------------------------|---------------------------------------|
| Track gauge             | 1435 mm                               |
| Number of cars per unit | 2 to 6 cars; basic version 3-car unit |
| Unit length             | 71.276                                |
| Max. operational speed  | 160 km/h                              |
| Power supply            | 1 Diesel motor (560 kW) in each car   |
| Unladen weight          | approx. 168.5 tonnes                  |
| Maximum axle load       | 18.5 tonnes                           |

---

An “Eco-Mode” programme has actually been initiated in-service for the Class 185 to maximise energy efficiency [50]. The Eco-Mode Phase 1 yielded significant savings, and in the next phase of the work the fuel economy is expected to improve further. The core concept of Eco-Mode is the selective use of the 3 engines [163]. The work presented in the current thesis is inspired by the selective operation concept and investigates this concept more thoroughly from a mathematical and optimisation viewpoint. The technique discussed in this chapter uses global optimisation, theoretically leading to improved fuel economy and optimal operation of the three engines.

## 7.3 Engine description

### 7.3.1 Engine efficiency map

The diesel engine is represented by a nonlinear static efficiency map which describes the instantaneous Brake Specific Fuel Consumption (BSFC) of the engine with different engine speeds and engine output power.

$$\zeta = f(P_e, \omega) \quad (7.1)$$

$$\eta = \zeta / P_e \quad (7.2)$$

$$P_e = T \omega \quad (7.3)$$

Where  $P_e$  is the engine output power,  $\zeta$  is the fuel rate in grams per second

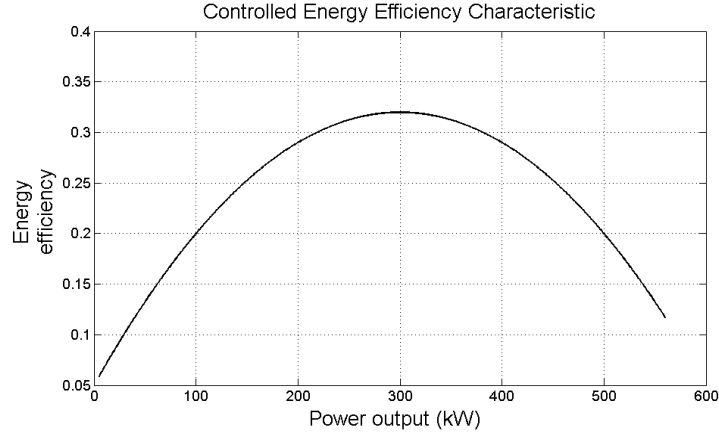


Figure 7.2: Controlled diesel engine power efficiency characteristic

( $gs^{-1}$ ),  $\eta$  is BSFC in grams per Joule ( $gJ^{-1}$ ),  $T$  is the engine output torque and  $\omega$  is the engine speed. To simplify, engine energy efficiency is defined by Eqn. 7.4:

$$\eta = P_e / P_f \quad (7.4)$$

Where the  $P_f$  is the power converted from the fuel at a fuel rate of  $\zeta$ .

The engine speed and power in this simulation have been limited to a single operating line on a typical engine efficiency map, i.e. each engine output power corresponds to a unique engine speed. The relationship between output power and fuel cost is convex [158]. This convex shape indicates that the most efficient operating point is below the maximum power output [151]. A characteristic fuel efficiency curve of a diesel engine used in similar simulation studies is displayed in Figure 7.2 [121]. In Figure 7.2, the engine is supposed to work along a controlled torque-speed trajectory in the efficiency map.

### 7.3.2 Diesel energy consumed calculation

Each iteration step in the STMS is one second, which is precise enough for the purpose of this study, thus the average power over that time step is numerically equivalent to the energy consumed in the time step. The total diesel energy consumption

is calculated by integrating the power history and considering the chemical conversion efficiency presented in Figure 7.2. Between one iteration step and the next, there is a limitation by the power slew rate of each engine. It is a limitation imposed on each engine to avoid any unrealistic output power jump, i.e. an engine cannot simply provide its maximum power output in one second. It has been assumed that in 7.3.1, within one second, the Maximum Power Switch (MPS)  $P_{mps}$  from one state to another is 30 kW. This means that the maximum total power difference from one second to the next is  $\pm 90$  kW. The value of  $P_{mps}$  can be changed but this does not affect the concept demonstrated in this study.

### 7.4 Problem definition

The optimisation is inspired by the fact that the total power demand can be met by a combination of unique power outputs from each engine. Additionally, the individual engine states are constrained by previous engine states and restrict the future engine states (governed by the positive and negative engine power slew rates). This enables a complex decision making procedure to be defined in order to calculate the required output power from each engine.

The first part of this section will introduce the optimisation objective and constraints. The second part will review the relationship between engine states and fuel cost.

#### 7.4.1 Objective and constraints

The objective of the power management strategy is to improve the fuel economy of the vehicle by distributing the power demand between the 3 engines. This can be formulated as an optimisation problem:

$$\min_x J(x) \text{ subject to } G(x) \leq 0 \quad (7.5)$$

---

#### 7.4. PROBLEM DEFINITION

---

The cost function  $J(x)$  represents the total fuel cost in an arbitrary duty cycle within a time length  $t_c$  as expressed in Eqn. 7.6.  $G(x) \leq 0$  accounts for the linear constraints shown in Eqns. 7.7.

$$J(P_A, P_B, P_C) = \int_0^{t_c} \text{fuelrate}(P_A, P_B, P_C) dt \quad (7.6)$$

The operating limitations of each engine set the boundaries for each engine's power  $P_A, P_B, P_C$ . Eqn. 7.7 illustrates that at any time during the journey, the engine cannot exceed its operating limit defined by Eqns. 7.8 and 7.9.

$$\begin{aligned} P_{A \min}(t) &\leq P_A(t) \leq P_{A \max}(t) \\ P_{B \min}(t) &\leq P_B(t) \leq P_{B \max}(t) \\ P_{C \min}(t) &\leq P_C(t) \leq P_{C \max}(t) \quad \forall t \in [0, t_c] \end{aligned} \quad (7.7)$$

The upper limit  $P_{e \min}$  and lower limit  $P_{e \max}$  of each engine's power output vary depending on the previous engine output  $P_{e \text{ pre}}$ , MPS (the maximum power slew rate)  $P_{\text{mps}}$  and Maximum Power Output. The decision process can be expressed as follows:

$$P_{e \min} = \begin{cases} P_{e \text{ pre}} - P_{\text{mps}} & \text{if } P_{e \text{ pre}} - P_{\text{mps}} > 0 \\ 0 & \text{if } P_{e \text{ pre}} - P_{\text{mps}} \leq 0 \end{cases} \quad (7.8)$$

$$P_{e \max} = \begin{cases} P_{e \text{ pre}} + P_{\text{mps}} & \text{if } P_{e \text{ pre}} + P_{\text{mps}} < P_{\max} \\ P_{\max} & \text{if } P_{e \text{ pre}} + P_{\text{mps}} \geq P_{\max} \end{cases} \quad (7.9)$$

It is assumed at every instance that the driver's power demand  $P_d$  should be satisfied since the driver's power demand should fall into the engine's operating limit. As a result, the following condition should be met:

$$P_d(t) = P_A(t) + P_B(t) + P_C(t) \quad (7.10)$$


---

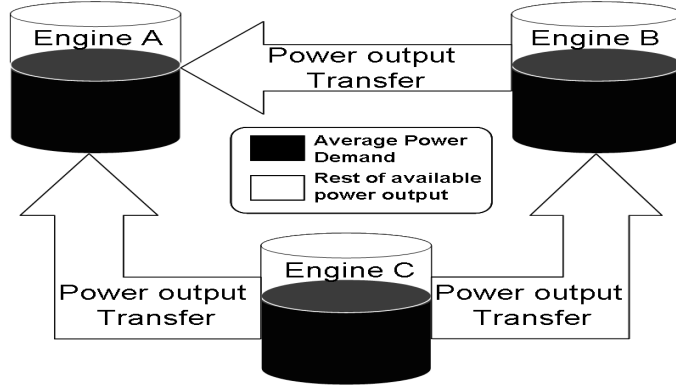


Figure 7.3: Engine power states approximation

### 7.4.2 Total power state vector and diesel fuel cost

The total power state vector is represented by  $[P_A(t), P_B(t), P_C(t)]$ . As stated in Eqn. 7.10, the entire engine output should meet the current power demand. If all the possible combinations of engine outputs are plotted in 3-D space, a triangle plane is produced as shown in Figure 7.4.

In Figure 7.4, some of the engine states, represented by black circles, have been plotted on a triangular plane. At any point on this plane the magnitude of the total power state vector is constant. Since the entire engine output must be non-negative, this triangular plane is only in the positive quadrant vector space, mathematically expressed as  $Q = ((X, Y, Z) | X, Y, Z \geq 0) \in R^3$ .

The number of possible engine power state vectors is innumerable. This will cause the search space for a minimum cost route to become too large to obtain an answer. The feasibility of dynamic programming strongly depends on the finite elements of the search space.

In order to produce a model with a finite number of engine power state vectors, some approximation has been made. This approximation makes the assumption that, within a selected power range, the change in energy consumption of an engine remains negligible. For example, if the power range is selected as  $[0, 30 \text{ kW}]$ , the difference in energy consumption for an engine output changing within the range of 30 kW is regarded as negligible.

---

#### 7.4. PROBLEM DEFINITION

---

An approximation procedure is presented in Figure 7.3. The procedure begins by dividing the power demand evenly into three portions, it then decreases the power output from Engine C, adding the amount of power reduction into Engines A and B. Considering all the possibilities for division between Engines A and B for each reduction in power output from Engine C, the procedure continues until all of the power output from Engine C has reduced to zero, or either of Engines A or B has increased to maximum power output. Engine B performs the same process by reducing its power output and adding the same reduction amount into Engine A. All along the process, the power output from Engine A is no less than Engine B, and Engine B is no less than Engine C, to avoid any vector duplication. The reduction amount for each engine is limited by the search gap previously defined.

With the assumption of the selected power range, the number of engine state vectors is significantly reduced. This approximation avoids missing any important power states in the searching procedure. For example, at a lower power demand, the searching procedure should be able to get access to the working mode with few engines.

An axis transformation is performed to demonstrate the relationship between the engine states and the fuel cost. The new Y axis and X axis are shown in Figure 7.4. The new Z axis is perpendicular to both the new X and Y axes and represents the fuel cost based on the data within the triangle.

The transformation has the following features:

1. The higher the power demand, the larger the area covered by the triangle;
2. The three vertices of the triangle each represent single engine operation;
3. The edges of the triangle represent dual engine operation;
4. Any point on the surface (not on an edge or vertex) represents triple engine operation.

After the axis transformation, by applying the fuel cost functions shown in



---

## 7.4. PROBLEM DEFINITION

---

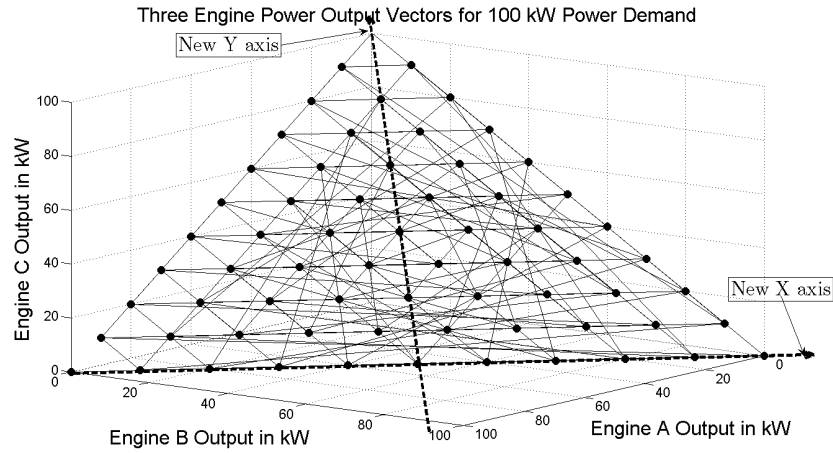
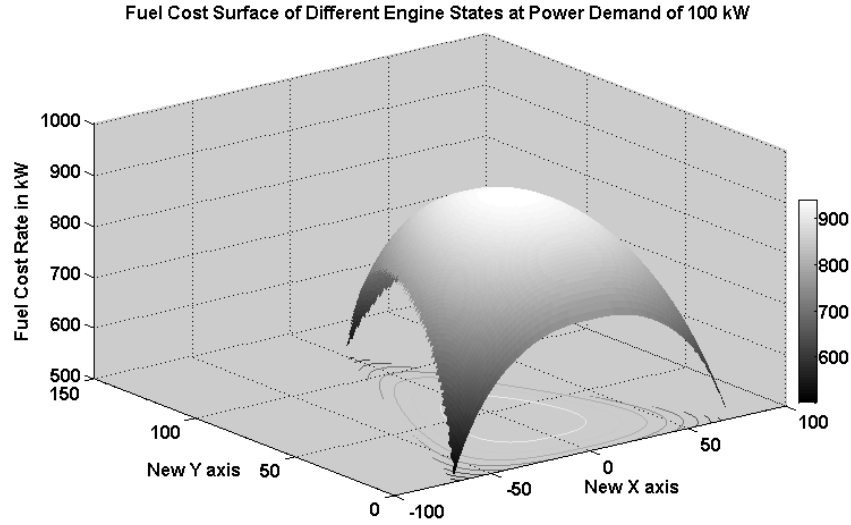


Figure 7.4: Total power state vector in 3-D spaces; the magnitude of the vector is 100 at any point at the triangular surface.

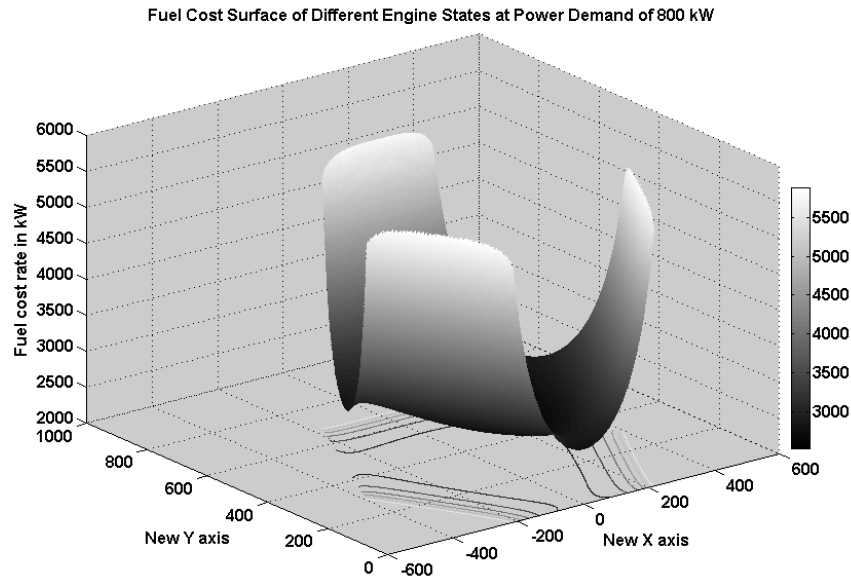
Eqns. 7.3 and efficiency characteristics shown in Figure. 7.2 described in Section 7.3, the relationship between different total power state vectors and fuel cost is demonstrated in Figure 7.4 for both high and low total power demands.

In Figure 7.5(a) and Figure 7.5(b), it is easy to identify that a significant change occurs when power demand increases from 100 kW to 800 kW. When the power demand is 100 kW, the fuel cost is highest in the center area and lowest in the edge area. This implies the most fuel efficient operation is to use one engine. At 800 kW, the shape of the fuel cost surface is inverted and the most efficient operation is where all three engines are used equally. It is noted that the rate of fuel can be transformed into rate of energy using the energy density data in [164].

Both of the cases in Figures 7.5 raise the possibility of complex dynamic decision making. Subject to the total power demand, the most efficient engine operation will change dynamically. The constraints of the power slew rate of each engine add additional complexity. The solutions and results of this problem are discussed in Section 7.5 using two different types of methodology.



(a) Fuel cost map for power demand of 100 kW



(b) Fuel cost map for power demand of 800 kW

Figure 7.5: Relationship between engine states and fuel cost under various power demands

## 7.5 Solutions and Results

This section proposes two typical methodologies for fuel efficient power management strategies [149]. Section 7.5.1 investigates the application of one global optimisation strategy: DP based on Bellman's Principle of Optimality [73]. Section

7.5.2 discusses an adaptive on-line power management strategy developed from DP. Due to the requirement for the whole journey profile to be known in advance, and the intensive CPU time associated with this method, DP is not suitable for on-line optimisation. However, the results from DP can provide expert system rules information which can be incorporated into on-line applications [165].

### 7.5.1 Dynamic Programming

#### Introduction

The general introduction of Dynamic Programming can be found in Section 3.3.

In this case study, the basic form of a dynamic programming functional equation is:

$$f(S) = \text{opt}_{d \in D(S)} \{R(S, d) \circ f(T(S, d))\} \quad (7.11)$$

where:

- $S$  is the engine state at each time step
- $d$  is the engine power change at each pair of adjacent time steps
- $R(S, d)$  is the fuel cost in the current engine state  $S$
- $f(T(S, d))$  is a next engine state transformation or transition function, assumed as a zero cost for engine states transition
- $\circ$  sums the fuel cost for a single decision.

#### Optimal Substructure

An optimal decision consists of a series of engine state vectors, i.e.  $[P_A(t), P_B(t), P_C(t)]$ , where  $P_A(t)$ ,  $P_B(t)$  and  $P_C(t)$  is the instant power output from engine A, B, and C respectively. The optimisation problem is now changed into seeking an optimal decision in the decision space which could minimise the total fuel cost defined in Eqn. 7.11.

To illustrate this idea more explicitly, a diagram is shown in Figure 7.6.

In Figure 7.6, each circle symbolises one possible engine state at that power demand. There are  $N$  power demand data and each power demand is denoted by  $P_i$ . Here  $i$  is the time sequence number of the power demand.  $S_{i,j}$  denotes the engine state vector for the  $j_{th}$  engine state at power demand  $P_i$ , where  $j \leq M_i$  if there are  $M_i$  possible power state vectors for a power demand of  $P_i$ . Each engine has a fuel cost based on the fuel efficiency map in Figure 7.2, denoted as  $E_{i,j}$  for each power state vector  $S_{i,j}$ . The solid line arrows represent the search directions and the dashed arrows are for the optimal path.

Generally, there are innumerable possible power state vectors. To simplify the search, some assumptions have been made:

- There are two types of state vector which should always be included in the engine state vector search space. The reason for that is because these two types of state vector represent two extreme working conditions of engines. One is an evenly operating mode vector, in which power demand is evenly divided between all the engines, and the other is the least engine operating mode vector, in which the fewest possible number of engines are used to achieve the power demand;
- The engine power states are assumed to have quantised values (with an increment of 30 kW). The value of the increment is nearly 1/20 of the maximum engine power output and this can offer reasonable search precise. This will reduce the number of possible power state vectors without compromising the grade of global optimality;
- To avoid duplication of power state vectors, one assumes that power output from engine A is no less than B and power output from B is no less than C. For one engine state vector, it makes no difference if one of the engine output powers is generated by engine A or B. The order of the elements in a engine state vector is not important and what it matters is the combination of the

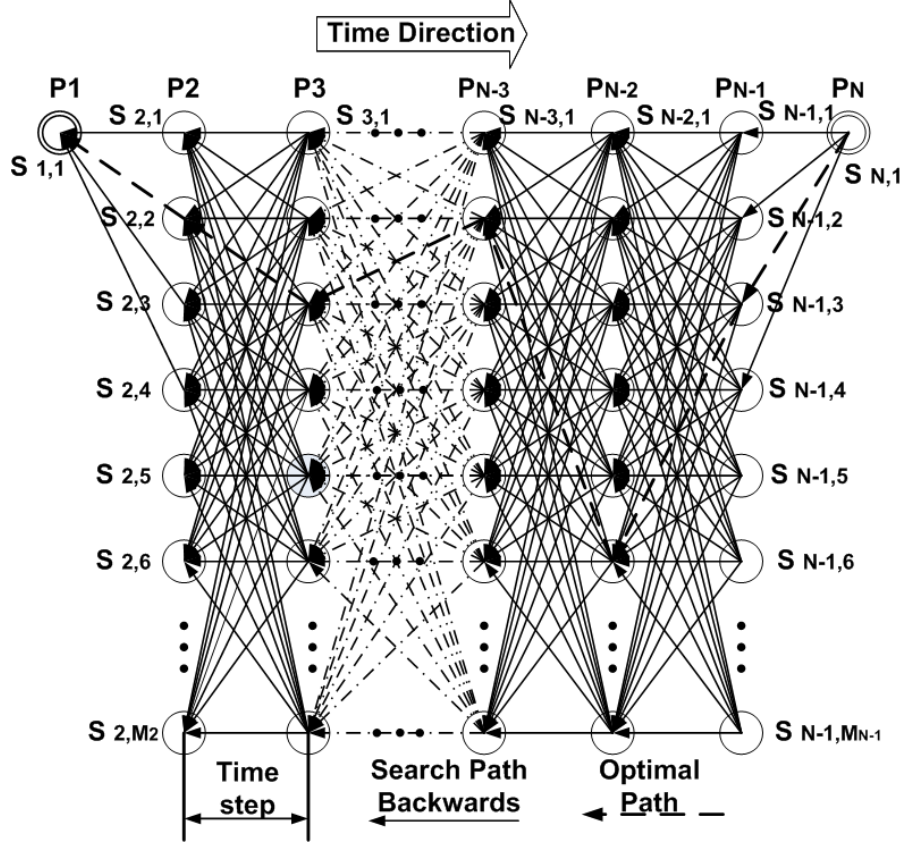


Figure 7.6: Dynamic programming evolution diagram

elements;

- During the braking and stopping operation, the power demand is zero and all engine outputs are zero. The engine does not need to provide any power. Since both the braking and stopping operation are necessary for all the solutions, the searching route will definitely go through zero output states, i.e. the zero power states are one of the optimum state series. The optimisation could be operated separately from a zero power demand to the next zero power demand in a smaller searching space, exactly as shown in Figure 7.6;

The optimisation procedure is performed backwards shown in Figure 7.6. Suppose that the minimum fuel cost from  $S_{N-3,2}$  to the final state  $S_N$ , is through  $S_{N-2,6}$ . Now it can be concluded that the fuel cost from states  $S_{N-2,6}$  to the final state  $S_N$  must also be the minimum. This is the case because if there are other possible routes between  $S_{N-2,6}$  and the final state  $S_N$  which cost less fuel and are also reach-

able from  $S_{N-3,2}$ , this path could easily be substituted to yield a lower fuel cost from  $S_{N-3,2}$  to the final state  $S_N$ . More generally speaking for such engine state problems, an optimal solution to a problem (searching for a minimum fuel cost through  $S_{N-3,2}$  to the final cost), contains an optimal solution to the sub-problem, i.e. obtaining the minimum fuel cost engine state vector series through  $S_{N-2,6}$ . This property is referred to as optimal substructure, as illustrated in Section 3.3.3, making DP a suitable method for such a problem [77].

### Solution

Dynamic Programming aims to find an optimal solution recursively from the optimal solutions to sub-problems as discussed in Section 7.5.1.  $F_i[j]$  denotes the minimum fuel cost engine path from engine states  $S_{i,j}$  to the final engine state. If  $F_i[j]$  finally defines the minimum fuel cost engine path from  $S_{1,1}$  to final states  $S_{N,1}$ , the optimal engine states path is found, denoted as  $F^*$ .  $F^*$  is defined in Eqn. 7.12:

$$F^* = \min(E_{1,1} + F_2[1], E_{1,1} + F_2[2], \dots, E_{1,1} + F_2[M_2]) \quad (7.12)$$

Where,  $E_{1,1}$  is the energy consumption of the first engine state. For the first step, from  $S_{N,1}$  to  $S_{N-1,j}$ ,  $F_{N-1}[j]$  could be defined in Eqn. 7.13:

$$\begin{aligned} F_{N-1}[1] &= E_{N,1} + E_{N-1,1} \\ F_{N-1}[2] &= E_{N,1} + E_{N-1,2} \\ &\dots \\ F_{N-1}[M_{N-1}] &= E_{N,1} + E_{N-1,M_{N-1}} \end{aligned} \quad (7.13)$$

Note that if  $S_{N,1}$  is not reachable for any engine states  $S_{N-1,j}$ , the corresponding  $F_{N-1}[j]$  should be set to infinity to rule this engine path out of the future searching range. Also note that, for each  $F_{N-1}[j]$ , the next engine state to achieve the minimum fuel cost  $F_{N-1}[j]$  should be stored to find the optimum engine states path. It

is proposed that another storage space is specified, named as  $R_i[j]$  to store the next engine state for  $F_i[j]$ .

Now consider the more general case for any  $F_i[j]$  where  $i = 1, 2, 3, \dots, N - 1$  and  $j = 1, 2, 3, \dots, M_i$ . The minimum fuel cost engine path through engine state  $S_{i,j}$  can be derived from the minimum fuel cost engine path from previously calculated minimum fuel cost  $F_{i-1}[j]$  where  $j = 1, 2, 3, \dots, M_i$ . The definition of  $F_i[j]$  can be found in Eqn. 7.14.

$$F_i[j] = \min(F_{i-1}[1] + E_{i,j}, F_{i-1}[2] + E_{i,j}, \dots, F_{i-1}[M_{i-1}] + E_{i,j}) \quad (7.14)$$

By combining Eqns. 7.13 and 7.14, a more general form for the minimum fuel cost  $F_i[j]$  is defined.

$$F_i[j] = \begin{cases} E_{N,j} + E_{N-1,1} & \text{if } i=N \\ \min(F_{i-1}[1] + E_{i,j}, F_{i-1}[2] + E_{i,j}, \\ \dots F_{i-1}[M_{i-1}] + E_{i,j}) & \text{if } i \leq N \end{cases} \quad (7.15)$$

After the final minimum fuel cost  $F_1[1]$  is found, it is straightforward to find the next engine state to achieve this fuel cost from  $R_1[1]$ . By doing this recursively, each engine state through the whole optimum path can be identified and the minimum fuel cost engine states path is found.

### 7.5.2 Adaptive online strategies

Dynamic programming is able to obtain the global optimal engine states path for the journey. However, the whole journey profile must be obtained in advance. This is obviously not possible for online power management [165]. This section will discuss the development of an adaptive online strategy for a potential real time operation using the results generated from DP.

Table 7.2: Threshold power demand value

| $P_{th1}$ | $P_{th2}$ |
|-----------|-----------|
| 382 kW    | 686 kW    |

### Threshold power demand

Based on the results of Dynamic Programming, it is found that with a different total power demand, there is a preferred number of engines which produce the best fuel economy, referred to as the Preferable Number of Engines (PNE). When the power demand is intermediate, e.g. during the course of cruising, the global optimisation path shows that the best way to supply the traction power is to use only 2 engines rather than 3. This type of characteristic can be used to develop a rule for an on-line intelligent power management strategy. From the current power demand, it is possible to calculate the preferable engine power split between the three engines.

Figure 7.7 illustrates the general idea of preferred engine number decision making based on current power demand. The threshold values are decided based on the following two presumptions:

- There are no engine states where two of the engine output powers are zero for power demand above  $P_{th1}$ ;
- There are no engine states where any of the engine output powers are zero for power demand above  $P_{th2}$ .

These two threshold power values  $P_{th1}$  and  $P_{th2}$  are listed in Table 7.2 using the results from the DP optimisation as demonstrated in details in Figure 7.11(a).

### Problem Formulation

After the preferable number of engines based on the two thresholds  $P_{th1}$  and  $P_{th2}$ , has been chosen, each preferable engine output can be determined, as shown in Table 7.3.



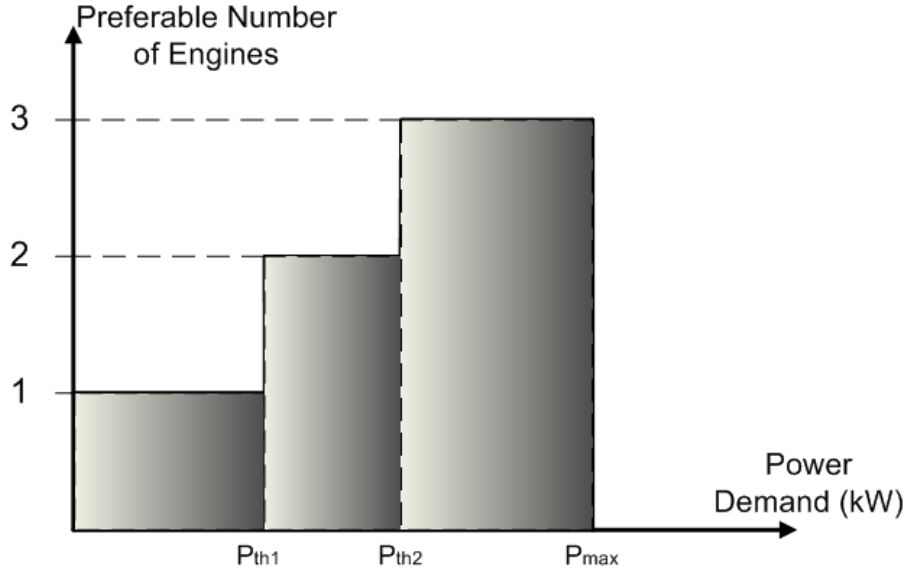


Figure 7.7: Preferable number of engines for various power demands

Table 7.3: Preferable power distribution based on various power demand

| Power Demand $P_d$ (kW)                                       | $NoE_{pre}$     | Preferable Power Distribution   |
|---|-----------------|---|
| $P_d(N\Delta t) < P_{th1}(N\Delta t)$                         | $NoE_{pre} = 1$ | $P_{pre}A(N\Delta t) = P_d$<br>$P_{pre}B(N\Delta t) = 0$<br>$P_{pre}C(N\Delta t) = 0$           |
| $P_{th1}(N\Delta t) < P_d(N\Delta t) \leq P_{th2}(N\Delta t)$ | $NoE_{pre} = 2$ | $P_{pre}A(N\Delta t) = P_d/2$<br>$P_{pre}B(N\Delta t) = P_d/2$<br>$P_{pre}C(N\Delta t) = 0$     |
| $P_{th2}(N\Delta t) < P_d(N\Delta t) \leq P_{max}$            | $NoE_{pre} = 3$ | $P_{pre}A(N\Delta t) = P_d/3$<br>$P_{pre}B(N\Delta t) = P_d/3$<br>$P_{pre}C(N\Delta t) = P_d/3$ |

$P_{pre}A(N\Delta t)$ ,  $P_{pre}B(N\Delta t)$  and  $P_{pre}C(N\Delta t)$  are the preferable power outputs for each engine. However, there is usually a discrepancy between current engine output and preferable engine output since the power demand changes dynamically. For example, at time  $n\Delta t$ , the total power requirement is satisfied with two engines working and the preferable engine output is equal to the current working engine output. At time  $(n + 1)\Delta t$ , the power demand increases, and engine C is brought into operation. Engine C was giving zero output power at time  $n\Delta t$ . For engine C, the discrepancy gap between current engine output and current preferable engine

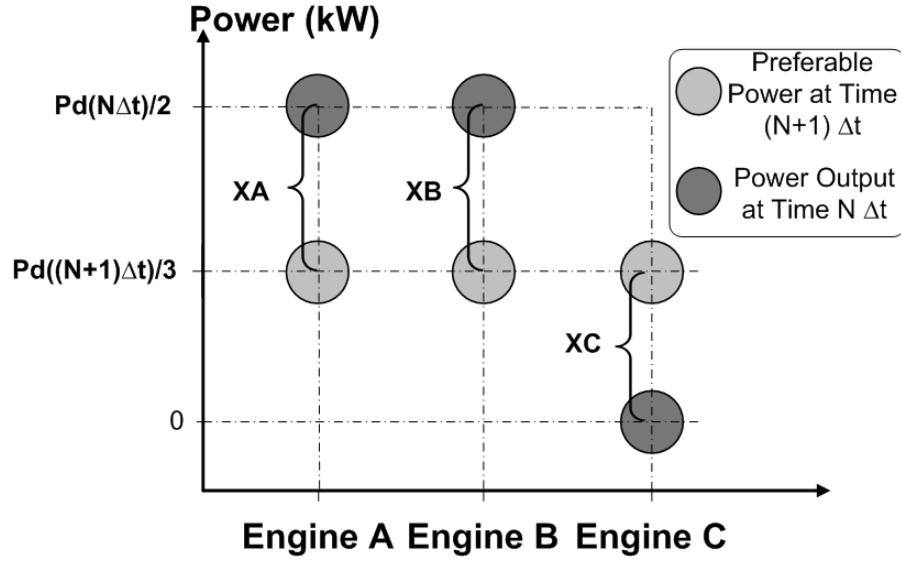


Figure 7.8: Illustration for minimisation of discrepancy

output could be very large.

Although it is not possible to achieve the current preferable engine output due to the power slew rate limit as illustrated in Section 7.3.2, it is possible to minimise the current total discrepancy to achieve an output power as near as possible to the preferable one using the non-linear optimisation techniques. Figure 7.8 illustrates the concept of minimisation of the discrepancy between the current engine output and the preferable engine output at the next time. At time  $N\Delta t$ , the power demand is  $P_d(N\Delta t)$ , and the preferable engine number is two. All the engines are working on their preferable states as shown by the grey circles, i.e.  $P_A = P_B = P(N\Delta t)/2$  kW and  $P_C = 0$  kW. At time  $(N+1)\Delta t$ , the power demand changes to  $P_d((N+1)\Delta t)$  and the preferable number of engines changes to three. At this time step, the preferable engine output changes as shown by the black circles in Figure 7.8. Each engine output is defined by  $P_{preA}((N+1)\Delta t)$ ,  $P_{preB}((N+1)\Delta t)$  and  $P_{preC}((N+1)\Delta t)$ , and  $X_A(N\Delta t)$ ,  $X_B(N\Delta t)$  and  $X_C(N\Delta t)$  denotes the discrepancy gap between the current engine output and the next preferable engine output.

Table 7.4: Online power distribution

|                        |   |
|------------------------|---|
| $NoE_{\text{opt}} = 1$ | $P_A((N+1)\Delta t) = P_d((N+1)\Delta t)$                       |
|                        | $P_B((N+1)\Delta t) = 0$  |
|                        | $P_C((N+1)\Delta t) = 0$  |
| $NoE_{\text{opt}} = 2$ | $P_A((N+1)\Delta t) = P_A(N\Delta t) + \Delta X_A^*(N\Delta t)$ |
|                        | $P_B((N+1)\Delta t) = P_B(N\Delta t) + \Delta X_B^*(N\Delta t)$ |
|                        | $P_C((N+1)\Delta t) = 0$  |
| $NoE_{\text{opt}} = 3$ | $P_A((N+1)\Delta t) = P_A(N\Delta t) + \Delta X_A^*(N\Delta t)$ |
|                        | $P_B((N+1)\Delta t) = P_B(N\Delta t) + \Delta X_B^*(N\Delta t)$ |
|                        | $P_C((N+1)\Delta t) = P_C(N\Delta t) + \Delta X_C^*(N\Delta t)$ |

$$X_A(N\Delta t) = P_A(N\Delta t) - P_{pref}A((N+1)\Delta t) \quad (7.16)$$

$$X_B(N\Delta t) = P_B(N\Delta t) - P_{pref}B((N+1)\Delta t) \quad (7.17)$$

$$X_C(N\Delta t) = P_C(N\Delta t) - P_{pref}C((N+1)\Delta t) \quad (7.18)$$

Here we define the total discrepancy gap at time  $N\Delta t$  denoted by  $D_t$  as follows:

$$D_t(N\Delta t) = |X_A(N\Delta t)| + |X_B(N\Delta t)| + |X_C(N\Delta t)| \quad (7.19)$$

The optimum change power of each engine will be the change that minimises  $D_t((N+1)\Delta t)$ .  $D_t^*((N+1)\Delta t)$  denotes the minimum total discrepancy,  $\Delta X_A^*(N\Delta t)$ ,  $\Delta X_B^*(N\Delta t)$ ,  $\Delta X_C^*(N\Delta t)$  is each change of engine output. The result is:

$$D_t^*((N+1)\Delta t) = \min(D_t((N+1)\Delta t)) \quad (7.20)$$

A set of variables  $\Delta X_A$ ,  $\Delta X_B$  and  $\Delta X_C$  are introduced to minimise  $D_t(N\Delta t)$ , where  $\Delta X_A$ ,  $\Delta X_B$  and  $\Delta X_C$  denotes the change of each engine power output during the next time step, shown in Eqn.7.24.

To find the set of engine power change values for the minimum total discrepancy, a constrained nonlinear optimisation has been implemented [166]. The Opti-

## 7.5. SOLUTIONS AND RESULTS

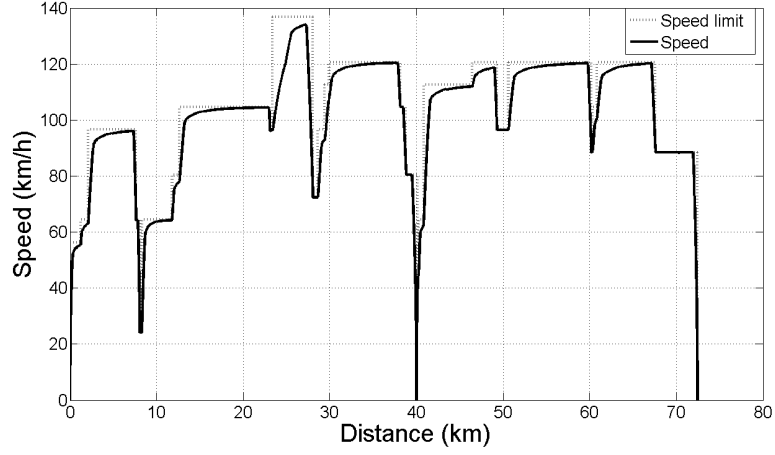


Figure 7.9: Speed profile and speed limit

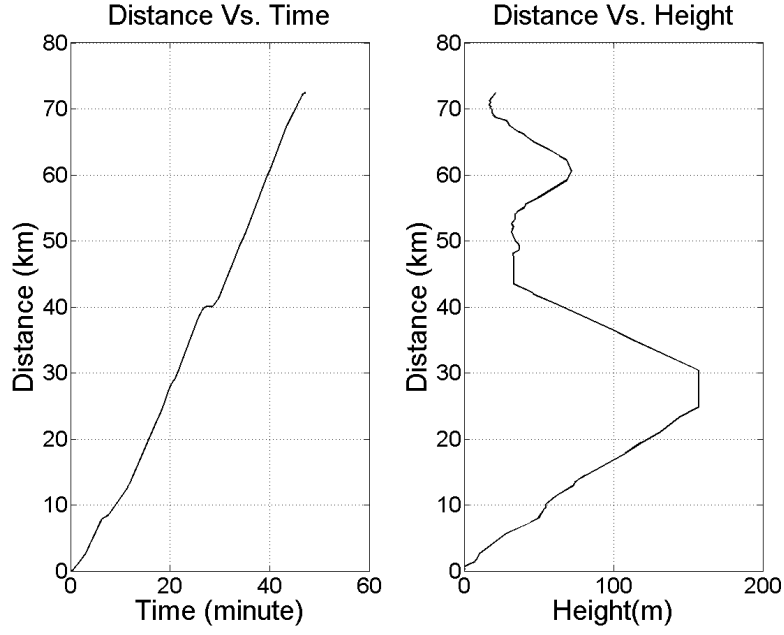


Figure 7.10: Distance time and distance elevation graph

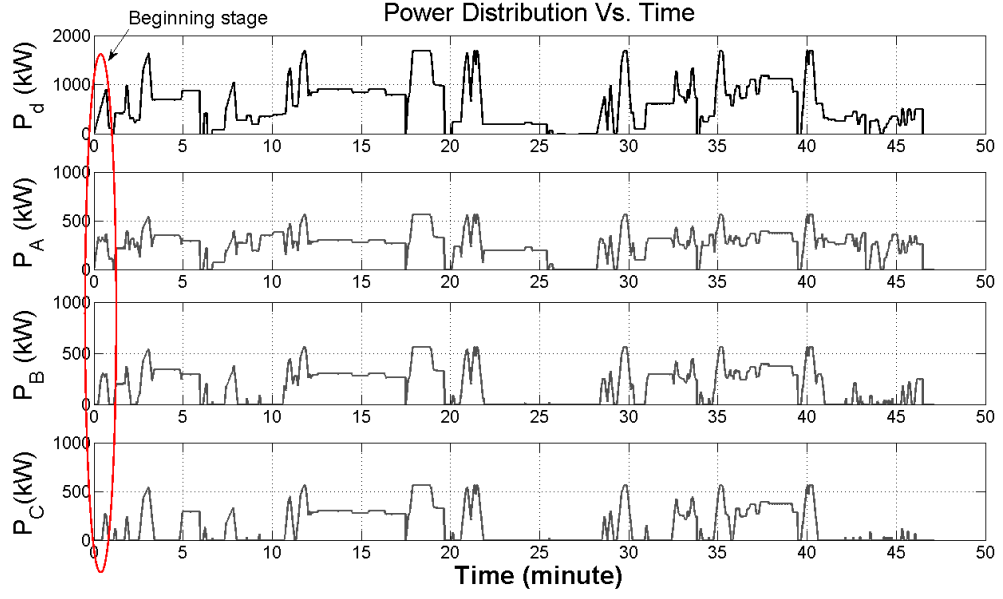
misation Toolbox<sup>TM</sup> in MATLAB is used to solve the problem. The constraints are listed below:

$$-P_{mps} \leq \Delta X_A(N\Delta t) \leq P_{mps} \quad (7.21)$$

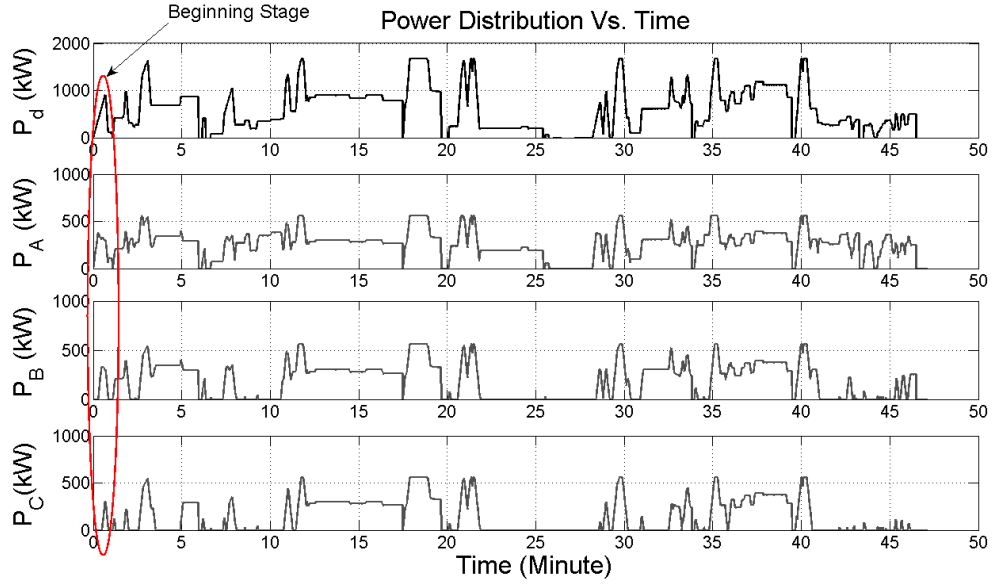
$$-P_{mps} \leq \Delta X_B(N\Delta t) \leq P_{mps} \quad (7.22)$$

$$-P_{mps} \leq \Delta X_C(N\Delta t) \leq P_{mps} \quad (7.23)$$

## 7.5. SOLUTIONS AND RESULTS



(a) Power distribution over time using dynamic programming



(b) Initial stage of power distribution by online rule based

Figure 7.11: Power distribution over time using dynamic programming and online rule based strategy (separate version)

$P_{mps}$  is the positive MPS in Section 7.3.

$$\Delta X_A(N\Delta t) + \Delta X_B(N\Delta t) + \Delta X_C(N\Delta t) = P_{diff}(N\Delta t) \quad (7.24)$$

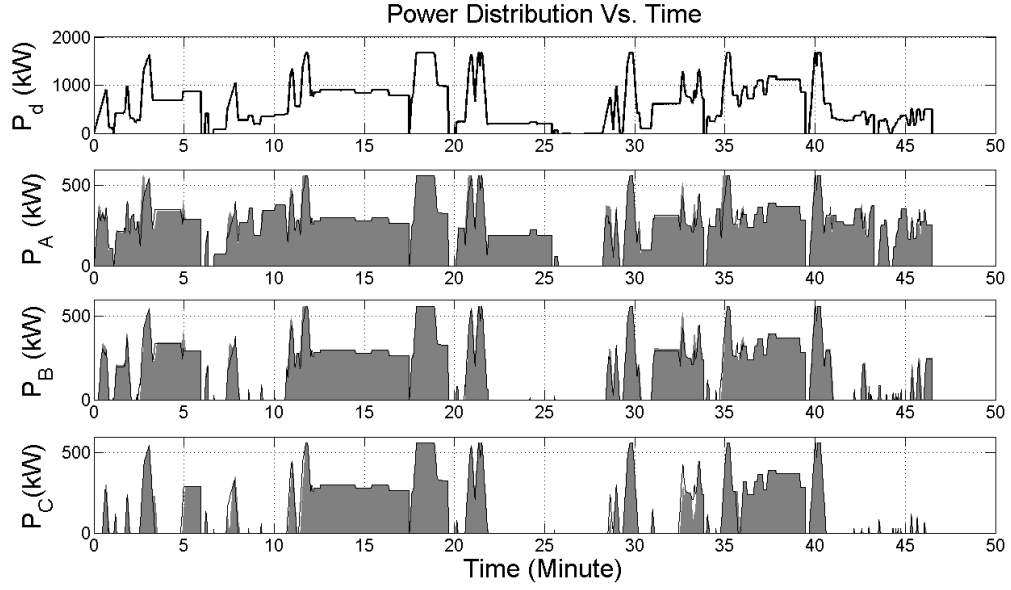


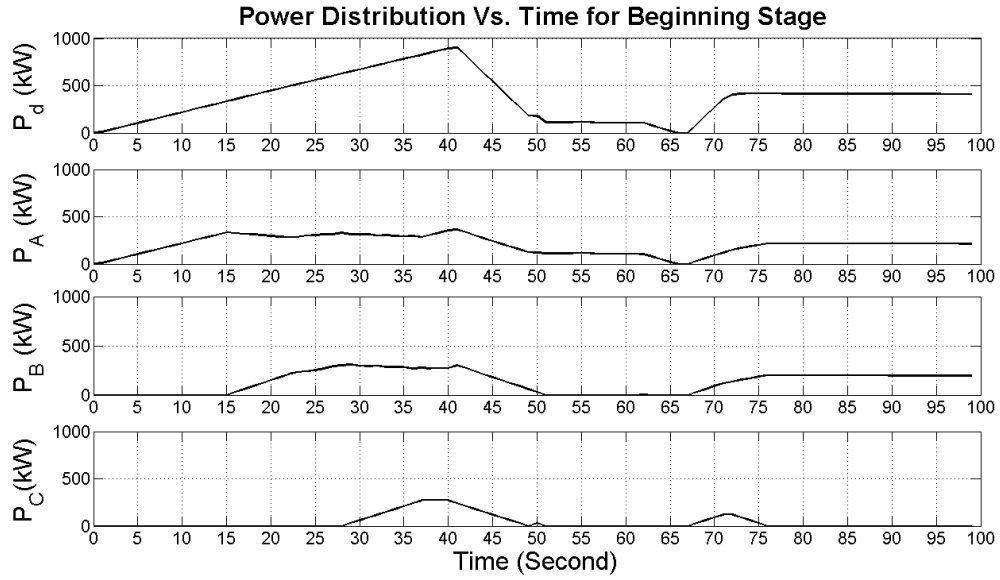
Figure 7.12: Power distribution over time using DP and online rule based strategy.

$$NoE_{opt} = \begin{cases} \max(NoE_{pre}, NoE_{act}) & \text{if } P_{diff}(N\Delta t) < 0 \\ \max(NoE_{pre}, NoE_{act}, NoE_{min}) & \text{if } P_{diff}(N\Delta t) \geq 0 \end{cases} \quad (7.25)$$

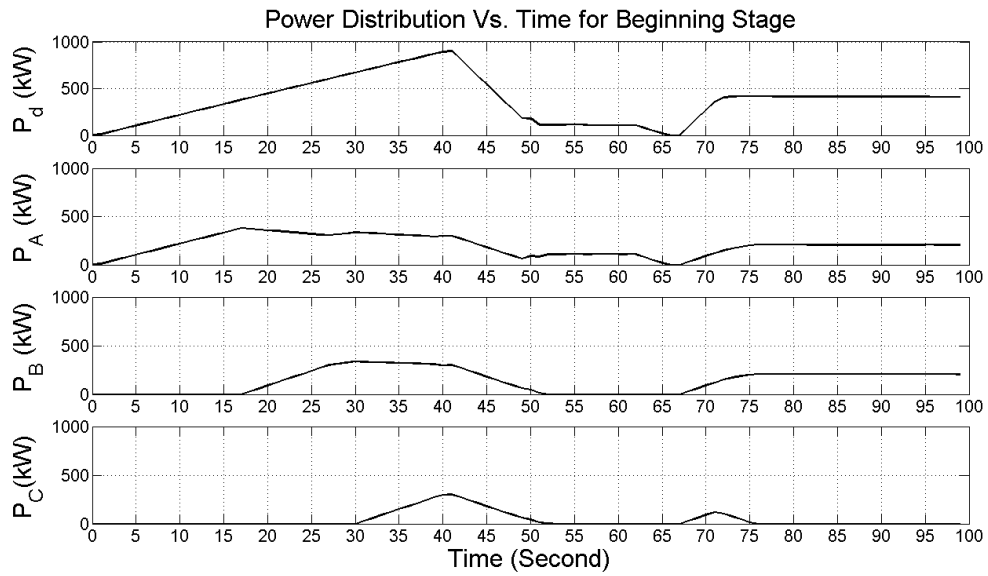
$P_{diff}(N\Delta t)$  is defined as  $P_d((N + 1)\Delta t) - P_d(N\Delta t)$  which is the power demand difference between the current total engine output and the next total engine demand. These constraints are set to ensure the total power demand can be met while the individual engine power changes are minimised. Some points should be noted to avoid undesirable results for practical operations. Some of the notations are listed below:

- $NoE_{pre}$  denotes the current preferable number of engines
- $NoE_{act}$  denotes the number of current working engines
- $NoE_{min}$  denotes the minimum number of working engines for the next power demand
- $NoE_{opt}$  denotes the number of engines for optimisation, i.e. the minimisation

## 7.5. SOLUTIONS AND RESULTS



(a) Initial stage of power distribution by DP



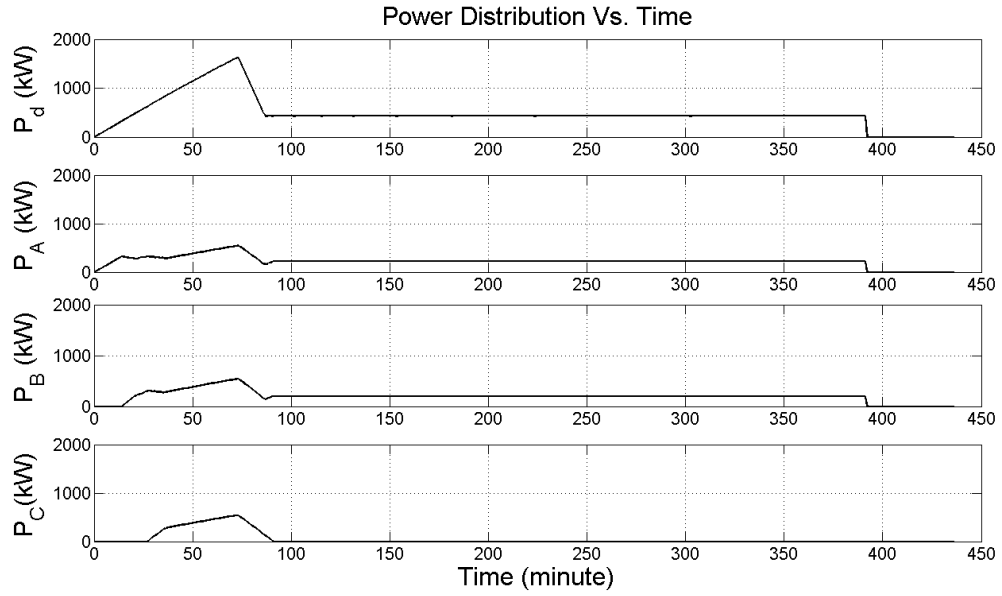
(b) Initial stage of power distribution by online rule based

Figure 7.13: Initial stage of the power distribution for both power management strategies

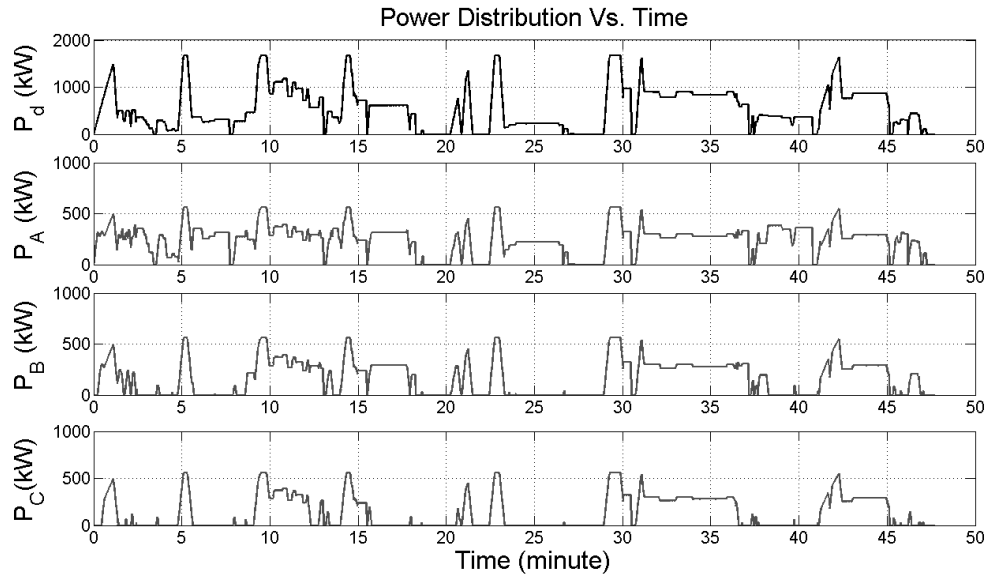
process is not always applied for all numbers of engines.

The online power distribution is summarised in Table. 7.4.

## 7.5. SOLUTIONS AND RESULTS



(a) Power distribution by DP for simple journey case study



(b) Power distribution by DP for a return journey case study

Figure 7.14: Power distribution by dynamic programming for two further case studies

### 7.5.3 Results

Figure 7.9 shows the speed profile over the route, while showing the speed limit at various sections of the journey. There is a station at 40 km.

The relationship between distance and time has been depicted in the left subplot of Figure 7.10. The right subplot, for comparative study purposes, displays the



distance vs. height so that both graphs can be read to explain the current height at a particular time instance.

The power distribution over the journey time is shown separately in Figure 7.11 and together in Figure 7.12. Since the negative braking power demand results in a zero power output from each engine, it is omitted from these figures.

Figure 7.13 shows that at the beginning of the journey, the power demand is increased from zero to a maximum point to accelerate the vehicle. During this course, each engine starts in sequence. This can be explained by the Preferable Engine Number computation, i.e. the preferable engine number for the best fuel economy varies during the initial acceleration phase. DP helps to locate the power change instant for each engine.

The other three subplots show the power output from engines A, B, and C respectively. Each of the subplots shows results from both of the proposed strategies. Figure 7.13 shows a magnified view at the beginning stage for both strategies. The high level of agreement is due to the intrinsic sub-optimal characteristics. In the rule-based online strategy, the preferable engine output is an approximation of the results from DP and not all of the approximation can match the global optimisation result. Nevertheless, the power distributions using both strategies depict a high degree of agreement. In Figure 7.12, the power split computed using the rule based strategy is shown by the solid grey area, and the power split computed using DP is shown by the solid black line. There is very little difference between the two methods, however some differences can be observed, for example the power demand of engine C at about 8 minutes, and after 33 minutes. To summarise, real time operation of power management can benefit from the online strategies developed using off-line global optimisation techniques.

In order to showcase a broader prospect of the application of dynamic programming, two extra case studies have also been done. In the simple case study of a 10 km journey, see Figure 7.14(a) for details, the vehicle will accelerate to 96 km/h with the flat gradient until braking and decelerate to the next station. The

Table 7.5: Simulation results comparison

| Method                 | Energy Consumed (kWh) | Fuel Cost (kg) |
|------------------------|-----------------------|----------------|
| Dynamic programming    | 1888.3                | 149.1          |
| Rule-based online      | 1918.4                | 151.4          |
| Empirical evenly split | 2019.8                | 159.5          |

second engine is switched on subsequently at the beginning of the journey and the third engine is kept off until the end of the journey due to the low power requirement from the driver. In the case study on the return journey, the vehicle is set to run backwards along the same route. The results for the power distribution are demonstrated in Figure 7.14(b). It is shown that dynamic programming is capable of calculating the distribution between the three engines in a different scenario.

Table 7.5 shows the simulation results for the three power management strategies. The empirical evenly split strategy means that the power requirement is evenly divided between 3 engines all the time. The conversion between the energy consumed and fuel cost is based on the energy density [164]. As suggested, the energy density for diesel oil for automotives is 45.6 GJ/tonne. The saving rates for both Dynamic Programming and Rule-based online strategies compared to an empirical evenly split strategy are 6.52% and 5.08%.

## 7.6 Summary

Recently, many power management strategies have been applied and realised for hybrid vehicle propulsion systems. These strategies are applied to achieve improved fuel economy and better environmental performance. However, so far, these technologies have not been considered for a conventional propulsion system. This chapter takes a typical British DMU railroad vehicle as a case study and explores the applicability and transferability of an optimal power management strategy. Due to the electrical inter-connection configuration, which enables fewer engines to work during the lower power requirement journey range, more advanced control

strategies are becoming applicable for power management systems.

This chapter firstly discussed a platform for the study of power management strategies: STMS. Through discretisation of the Newton equations and modern resistance calculation (*Davis Equation*), a STMS is modelled. This simulator is able to take account of the geography and driving style to produce an online output of the train position, speed, and current power usage, etc. Based on this, dynamic programming, together with an adaptive on-line rule-based strategy for this typical British DMU train are proposed in this chapter. Some improvement and modification has been adopted to decrease the search complexity and thus total computation time while maintaining acceptable optimisation precision. Based on the results of dynamic programming, an adaptive online rule-based strategy has also been discussed for further online simulation using large scale non-linear programming.

After transferring the energy consumed into diesel fuel cost, both of the simulation results indicate that a remarkable energy saving has been achieved using both of the two strategies.

This chapter considers the engine efficiency map in two dimensioned space. Further investigation is needed for the study of transient characteristics of the diesel engine.

# Chapter 8

## Conclusions and future work

### 8.1 General summary of contents

The contents of this thesis can be divided into two parts.

- Part one presents the background and literature review which includes the first three chapters. Chapter 1 introduces different energy efficiency improvement techniques which have been implemented in industry and trialled in other studies. Rather than developing a new set of electric devices or control techniques, using a high-level supervisory operational control philosophy will incur much less cost, and will also increase the energy efficiency significantly. Chapter 2 is a review of railway traction systems including from traction supply types as well as the electric drive and both AC and DC machines. Chapter 3 reviews mathematical optimisation techniques. The discussion in this chapter is mainly divided into numerical optimisation and metaheuristics. DP is introduced separately due to its significance.
- Part two of this thesis describes the application of the proposed high level supervisory control in various scenarios. An introduction to modelling train motion and traction power is presented. Three case studies under the pre-set research scenario are discussed from Chapter 5 to Chapter 6. In Chapter

5, based on the proposed distance based searching model for train running trajectories, various optimisation techniques: DP, ACO and GA are applied to search for the optimal energy efficient train speed trajectory under certain time consumption constraints. In Chapter 6, a railway DC electrical network is modelled and simulated using both NA and LF calculation. Based on this model, the instant electrical states of multiple trains running on two parallel railways can be obtained while considering the characteristics of the railway traction power, the DC substations and the electric power transmission circuit. In Chapter 7, a case study of a typical British DMU train is conducted to investigate the application potential of advanced energy management strategies. It is found that by using a set of energy management strategies, the distribution of power demand can be intelligently performed so that equipped engines can operate in a higher efficiency area on the efficiency map. The total fuel cost is found to be significantly improved .

## **8.2 Conclusions**

The major findings and contributions are concluded as follows.

### **8.2.1 Single train trajectory optimisation study**

Energy efficient driving strategies for railway vehicles are becoming even more important. Under certain operational, geographic and physical constraints, the energy efficiency for the train speed trajectory can be significantly improved if effective train control strategies are used. Due to the nonlinear constraints, the solution for using control strategies is not easily obtained. There are two types of strategy applied to obtain the energy efficiency train speed trajectory. One of them is to effectively apply coasting operations to obtain a compromise between the running time and energy cost. The other is to search for the feasible speed control strategy to ultimately reduce energy consumption while maintaining the journey time require-

---

## 8.2. CONCLUSIONS

---

ment. This thesis proposes a distance based train trajectory search model, upon which three algorithms have been applied to search for the optimal train speed trajectory. Instead of searching for a detailed, complicated control input for the train traction system, this model tries to obtain the speed level at each preset position along the journey. Based on the results depicted in Section 5.7, the major findings in this study are concluded as follows.

- By arbitrarily creating connected searching graphs consists of vehicle states, Ant Colony Optimisation shown in Section 5.4, Genetic Algorithms in Section 5.5 and Dynamic Programming in Section 5.6 are demonstrated to find energy efficiency train trajectory trajectory.
- The performance of 3 methods have been contrasted and compared. It was found that Dynamic Programming performed better than both the Genetic Algorithm and Ant Colony Optimisation. Under certain circumstances the Genetic Algorithm performed quite poorly and failed to converge onto a good solution (particularly for large journey times). It may be possible to tune the search algorithm, but without comparative results from alternative methods it would be impossible to determine the existence of better solutions. Ant Colony Optimisation depended on strong heuristic information and performed adequately well for most of the journey time allowances. It also arrived at a solution significantly quicker than the other two methods.
- For those cases where the solution space becomes small, both the Genetic algorithm and Ant Colony Optimisation failed to converge on a solution. When the time requirement is even less, it is required to reach plausible train trajectory using more rigorous selection of candidate velocity at different distance. However, the algorithms based on the random Monte Carlo style selection shows less capability to reach optimum trajectory in a limited number of iteration steps.
- Dynamic Programming is the most recommended algorithms to search for

the energy efficiency train trajectory for an off-line application. Though the computational time is significantly increased, the results obtained by Dynamic programming are robust and resilient.

- More heuristic information improves robustness of algorithm. For example Ant Colony Optimisation will maintain its searching results quality when the searching spaces are increased due to its strong heuristic information. Genetic Algorithm loses its optimality due to less heuristic information involved.
- In general, it is recommended that more than 1 method should be used to identify optimum trajectories because it is often possible to converge on a solution which is plausible, yet nowhere near optimal.

### 8.2.2 Journey time allocation study

It is noted that in a DC railway network, power peaks are undesirable for the sake of system safety and energy efficiency. Several methods such as intelligent traction control method, synchronism management method etc. are proposed in literatures and coordinated operation between trains in a network is one of the more commonly used methods. This study adopted journey time allocation for each inter-station operation as the coordinating tool for trains. Furthermore, this study also considers journey time as an optimisation objective as well as the energy efficiency of the electrical network.

By varying the journey time allocation, the regenerative power utilisation ratio and energy loss in the substation can both vary because the synchronism of operations of trains may vary. After achieving train running trajectories under dynamic journey time constraints, in Section 6.3, this thesis further develops a multiple electrical train model with regenerative braking in a DC electrical network using the NA and LF methods. It demonstrates the relationship between the inter-station journey time allocation and the total journey quality in terms of journey time and energy

efficiency. In Section 6.3.3, a GA is applied to search the optimal journey allocations and the result is bench-marked by the BF method in Section 6.4. The major findings are concluded as follows:

- Using Nodal Analysis and Load Flow methods, it is able to model multiple trains within a DC railway electrical network. Regenerative braking and dynamic performance of electrical network can be simulated. See Section 6.2 for further details.
- Energy efficiency of the electrical network can be improved by reducing the power peaks in the substations which can be realised using operational methods such as distribution of inter-station journey time.
- Increasing part of inter-station journey time is one of feasible solutions to reduce power peaks and improve energy efficiency.
- Genetic Algorithm is a solution to find the energy-efficient distribution of inter-station journey time.

### **8.2.3 A power management strategy for a multiple unit train**

Improved fuel consumption and lower emissions are two of the key objectives for future transportation. HEVs, in which two or more power systems are combined, are able to meet these objectives through the capture and reuse of regenerated braking energy, and through optimised use of the prime mover. However, more complicated power management strategies are required for such vehicles. This thesis explores the potential of applying advanced power management strategies to a DMU train described in Section 7.2. These types of vehicles have multiple diesel engines, which are commonly operated in a homogenous manner. The work described in this thesis analyses the potential energy savings that may be obtained through the independent operation of the engines. Two widely investigated power management



---

## 8.2. CONCLUSIONS

---

strategies illustrated and have been applied to a typical multiple decisions making problem for a typical DMU train defined in Section 7.4.

In Section 7.5.1, DP strategies have been applied to the results produced by a STMS (see Section 4.3.4 for details) to identify the optimal instantaneous power distribution between the engines. In Section 7.5.2, an adaptive rule-based online strategy based on the optimisation results from the DP solution is then proposed and realised using a non-linear programming optimisation algorithm. Both of these strategies indicate acceptable agreement and show around a 7% fuel cost reduction in comparison with the evenly-split engine operation [49]. Based on the results demonstrated in Section 7.5.3, the major findings are concluded as follows:

- Advanced power management strategies can be applied in a conventional Diesel Multiple Unit train to reduce fuel cost.
- Dynamic Programming is applied to identify the optimal instantaneous power distribution between the engines. An adaptive rule-based online strategy based on the optimisation results from the Dynamic Programming solution is then proposed and realised using a non-linear programming optimisation algorithm.
- Both of these two strategies indicate acceptable agreement and Dynamic Programming shows around a 7% fuel cost reduction and Rule-based strategy shows around a 5% fuel cost reduction rate in comparison with the evenly-split engine operation.
- In order to save fuels, it is suggested that when a train should start up each engines sequentially when it sets off from a station. It is because fewer engines are required for low power demands.
- The optimised power management strategies become similar to a evenly-split strategies when the power demand is high when the train is running at high speed.

## **8.3 Future work**

### **8.3.1 Validation and verification**

All the studies covered in this thesis are based on theoretical simulation. It provides useful information for the energy efficiency development of railway traction engineering. Validation and verification remains to be the future work for the proposed study context. Without doubt, the demonstrated work in this thesis has proposed several novel energy-saving ideas. It is hoped that these ideas can provide aspirations for the further improvement of energy efficiency in the railway traction system.

### **8.3.2 Single train trajectory optimisation study**

The train trajectory between two pre-determined distance points is assumed to be fixed and a simplified train running pattern is applied as described in Section 5.3. As a result, some unrealistic trajectories may occur due to the simplicity of the train running pattern, for example, the train may have to brake before it reaches its maximum cruising speed. In the context of future research, it is worth searching the most efficient train speed trajectory between two predetermined distance positions using quick numerical searching methods.

It may also be interesting to consider trajectory optimisation for multiple trains in a network. As the train trajectory is also affected by other trains in a railway network, it is certainly of interest that the optimisation objective function can contain the total cost of multiple train trajectories. Such a study would be beneficial for the design of a signalling system.

### **8.3.3 Journey time allocation study**

For the sake of simplicity, the leakage current is omitted from the DC electrical network modelling in the journey time allocation study reported in this thesis. Mean-

while, the study of multiple objective optimisation between the journey time and energy efficiency needs further investigation. For multiple trains in an electrical network, the operation without considering the signalling system can be regarded as hypothetical. In addition, an unchanged headway time between subsequent trains is only used for research purposes. In summary, the work on journey time allocation reported in this thesis is worth more study from an engineering operations standpoint. In the next-stage research, more detailed and practical DC electrical network modelling can be interesting. Signalling system and flexible headway train running remain interesting for a future research context.

#### **8.3.4 Power management strategies study for a multiple unit train**

In the study of power management strategies for a multiple unit train, the energy efficiency characteristic shown in Figure 7.2 is assumed to be 2-D, in which energy efficiency only depends on the instant power output. However, in reality the power efficiency of a diesel engine is relevant to both torque and rotation speed. A more detailed investigation should be conducted before such energy distribution strategies are applied to a real multiple unit train traction system.

Further research should consider some other constraints. For example, more number of engines or electric motors can be installed to further increase the search space. Additionally, the application of such strategies should not be limited only to this field but can be expanded into any field with multiple power distribution demands. In some cases, the power demands are not known in advance and this will require a new method or improved Dynamic Programming.

On-line strategies based on the off-line global optimum strategies can be improved by using a less determinant rule-based system. Fuzzy logic and ANN are the candidate methods to improve the currently developed rule-based system.

# **Appendix A**

## **Electric drive for railway traction systems**

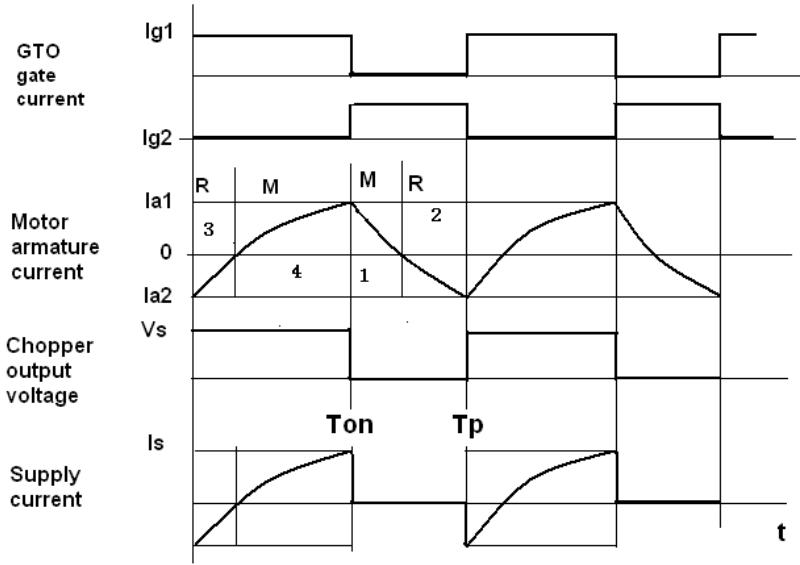
### **A.1 DC motor drive**

According to the method of power supply to the motor drives, there are two main types of DC traction drive: the chopper converter and the phase-control rectifier.

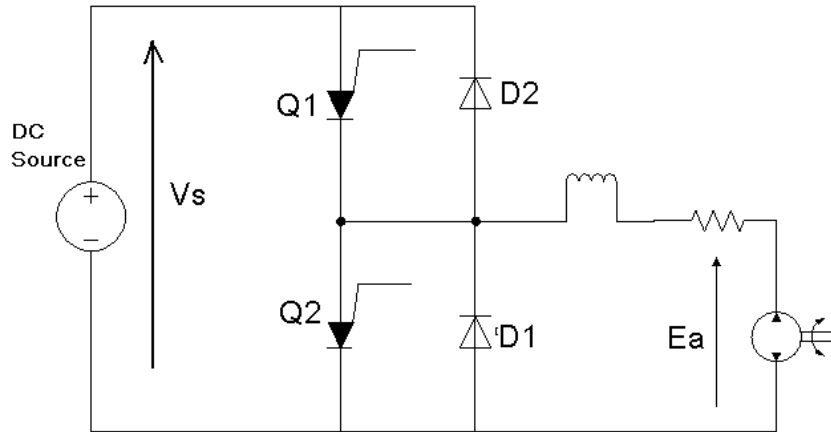
#### **A.1.1 DC-DC chopper converter traction drive**

The DC-DC chopper converter converts fixed DC voltage into various output voltages. Bi-directional operation is possible with a combination of step up and step down converters, which realises regenerative braking for railway vehicles. This type of converter is referred to as a two-quadrant converter for two different characteristics of operation. A DC-DC chopper converter helps to realise a smooth, step-less control and fast response to the target due to the flexibility of output voltage control. Additionally, compared to traditional electro-mechanical equipment, chopper operation significantly lowers maintenance costs as no mechanical requirement is needed in such a system.

The possibility and capability of regeneration of chopper converter traction led to the development of wayside storage substations. These are used as an alternative



(a) Schematic operational waveform



(b) Circuit diagram

Figure A.1: Two-quadrant traction chopper power operational circuit and the operational waveform

to the conventional way of consuming the energy recovered from nearby trains in new railways of Japan [62].

The thyristor  $Q_1$  turns on and off with firing angle  $\alpha$ . The current through the inductor increases during the “on” period as the voltage from the sources is higher than the terminal voltage on the motor and decays due to the back EMF applied. Though the current increases and then decays, the power flows from the power supply to the motor. The equivalent circuits during states 1 and 4 labeled in Figure A.1(a) are shown in Figure A.2.

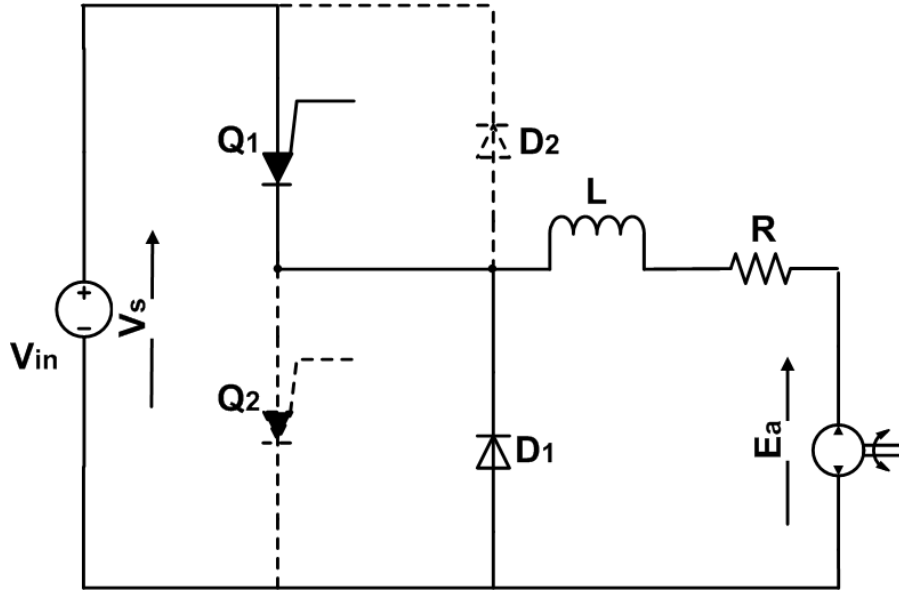


Figure A.2: Chopper power circuit corresponding to state 1 and 4 in Figure A.1(a)

With the current turning from positive to negative, the system changes into regenerative mode where the motor acts as a generator and energy is returned to the supply. At this stage, the operational devices are  $Q_2$  and  $D_2$ . With  $Q_2$  turning on, the current goes through and energy is stored in inductance  $L$ ; with  $Q_2$  off, the current decays due to the negative voltage difference  $V_s - E_a$ <sup>1</sup> applied on it and the power is delivered to the source via the diode  $D_2$ .

Figure A.3 illustrates the states of these operations.

One important feature of this circuit is that the thyristors do not work simultaneously to avoid the short circuit of the power supply. Two groups of devices,  $Q_1$ ,  $D_1$  and  $Q_2$ ,  $D_2$  work separately to realise the respective working mode. The speed or torque requirement determines the working mode of this chopper.

It is noted that the chopper operation inherently gives rise to the harmonics content to the output current and voltage [167]. The fundamental element of the harmonics caused by the chopper is the chopping frequency. In addition, the substation rectifier which rectifies AC power supply into the required DC power will usually introduce harmonics. Neither of these harmonics is desirable, due to their

<sup>1</sup>The source voltage  $V_s$  is higher than generator sources  $E_a$ .

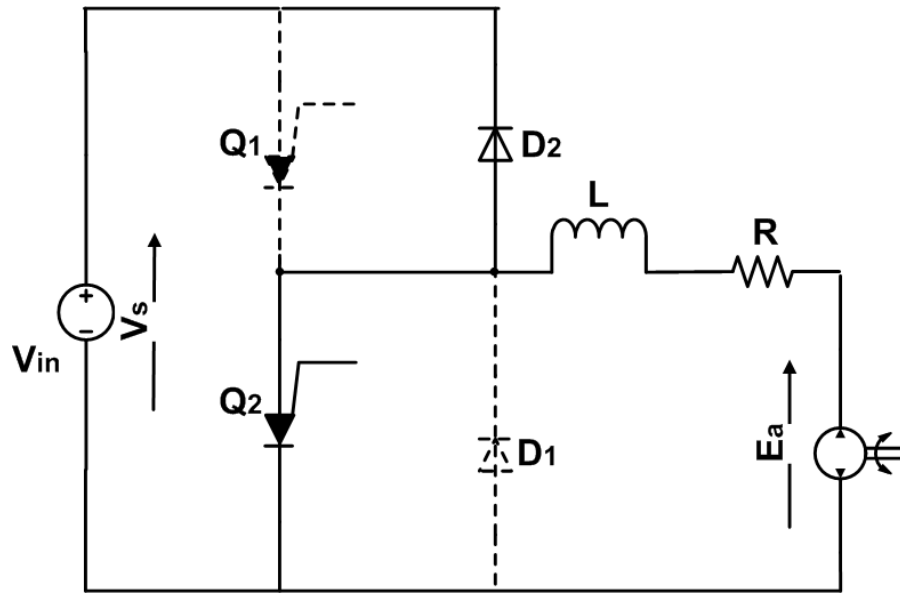


Figure A.3: Chopper power circuit corresponding to state 3 and 2 in Figure A.1(a)

interference with the signalling system along the railway line. Consequently, bulky filters are usually employed at the power collection terminal and within the chopper controllers to reduce the harmonics.

### A.1.2 Phase-controlled rectifier traction drive

The DC power supply may either be rectified from AC at the substation, as mentioned in the previous section, or by using an on-board rectifier, giving rise to phase-controlled rectifier traction drives.

It is preferable, as stated in [54], to adopt the half controlled bridge rectifier. The advantages of the half controlled bridge rectifier, as compared to the full controlled counterpart, are listed as follows:

1. Half of the thyristors are needed and the circuit costs less;
2. It operates more readily with a continuous current and leads to a higher power factor;
3. The armature current does not need to deal with negative voltage. The current will decay with commutating diode shown in A.4 and less torque ripple is

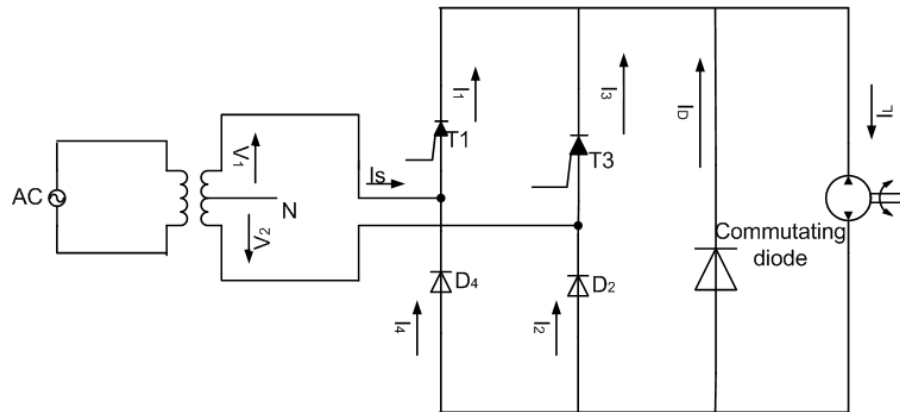


Figure A.4: Semi-controlled single-phase bridge operation circuit and waveforms

induced.

However, the AC supply line current will be more distorted and the inversion operation becomes more difficult.

The circuit diagram is shown in Figure A.4. Diodes  $D_2$  and  $D_4$  form a free-wheel path through which the current is drawn back in to AC source. In one period, the revolution could be divided into 4 divisions.

Before  $T_1$  is fired, the load current loops between the load. The commutating diode and the value of the current decays gradually as the firing angle increases to prolong the period of this current decaying process. During this interval, the voltage imposed on the load is zero.

At the firing angle  $\alpha$ , thyristor  $T_1$  is fired and the current goes through it, resulting in the voltage, being imposed on the load from the source. The current leads through  $T_1$ , the motor armature, and finally diode  $D_2$ .

As the diode will cut the current when the voltage through it becomes negative corresponding to the current direction, the current on the diode and thyristor becomes zero and the commutating diode is effective again until the next thyristor  $T_3$  is fired.

After thyristor  $T_3$  is fired, the load circuit works under the same conditions as the interval between  $\alpha$  and  $\pi$ . It is important to note the negative supply current



from the source.

## A.2 AC motor drive

### A.2.1 DC-fed Current Source Inverter drive

The current source inverter has the current from a DC source which is effectively constant over a period of a few cycles [167]. In practice, the constant current is realised by a large inductance and the value of the current is adjusted by the duty cycle of the pre-converter before the current source inverter, as shown in Figure A.6.

The devices in constant current inverters are required to withstand reverse voltages, due to the natural commutation; as a result, thyristors, rather than transistors should be adopted as the switching devices, unless an anti-parallel diode is included. For the purpose of circuit protection, relatively slow response thyristors are preferable, since it takes longer for the fault current to rise to a dangerous level and allows protection to be applied more effectively.

Figure A.6 shows the circuit diagram for a 3-phase constant current source inverter. Thyristors are fired in sequence to produce a quasi-square wave load current on each phase of the induction motor.

During regeneration, the current direction remains the same while the DC source voltage reverses. The current lags the voltage by less than  $\pi/2$ , however in generation mode this angle changes to the range between  $\pi/2$  and  $\pi$ , bringing about the leading angle between current and voltage.

Pulse Width Modulation (PWM), which replaces the square wave by a set of discrete pulses in the same duty cycle, could be applied to the output current of the inverter. The advantage of PWM is to eliminate the low order harmonics and achieve a more stable operation of the motor.

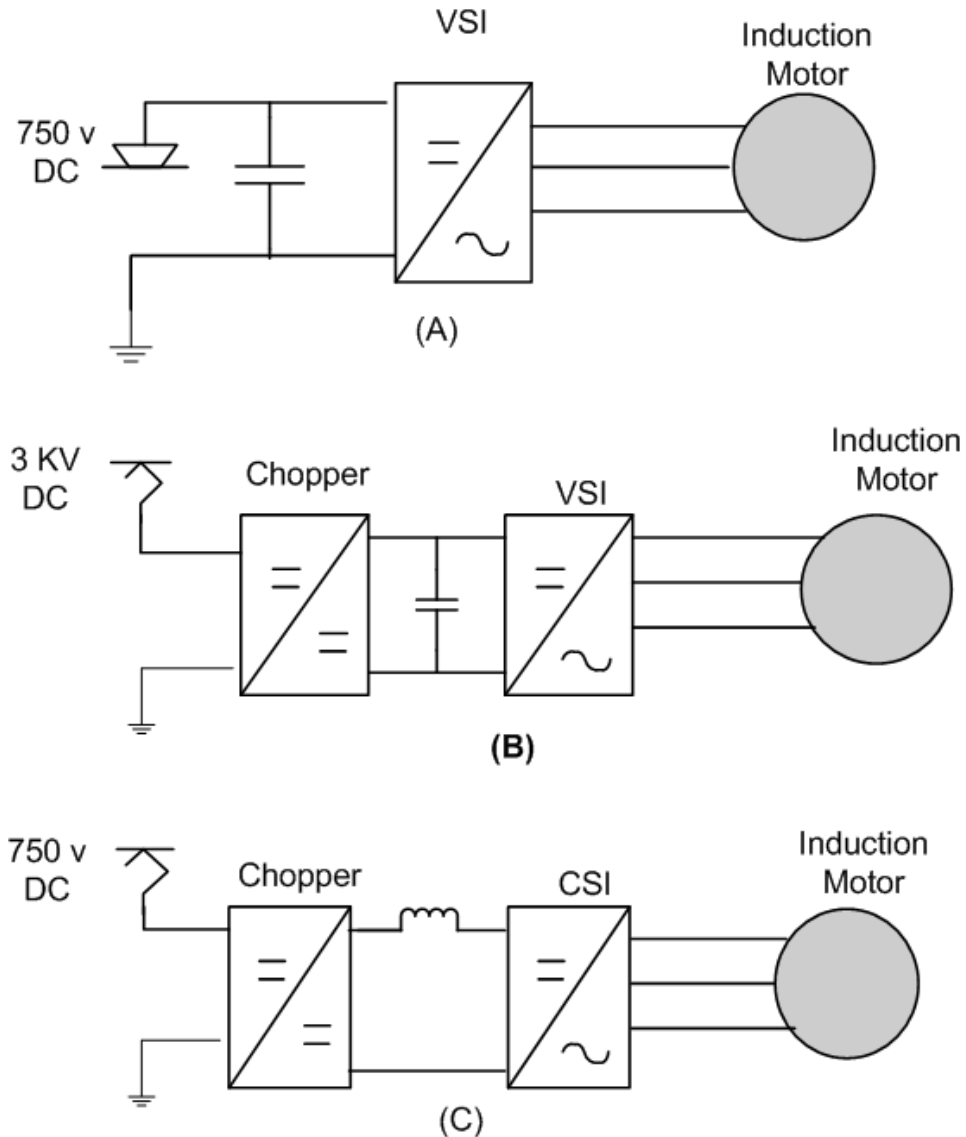


Figure A.5: Traction drives with three-phase motors and DC power supply

### A.2.2 DC-fed Voltage Source Inverter drive

The VSI inverts the constant voltage into various voltage outputs on the load. Constant voltage implies that over the short time of one cycle of the output AC waveform, any change in the DC source voltage is negligible.

The reversal of power flow occurs during braking. At this stage, the induction motor works as a generator and the inverter output frequency is lower than the rotor rotational speed. The current direction will reverse while the phase of voltage remains unchanged.

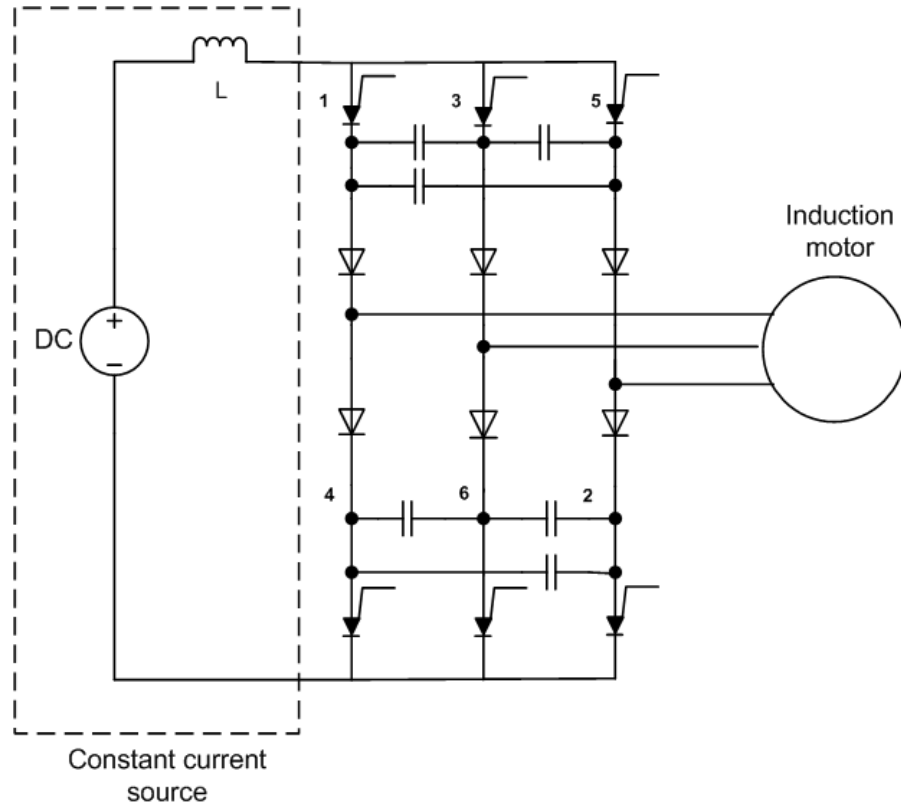


Figure A.6: Current source inverter circuit diagram

VSI Induction Motor (IM) traction drives for DC railways take two different types of inverters, two-level and three-level [53].

The two-level inverter limits its application to under 1.5 kV DC traction systems as the limited blocking voltage capability is 4.5 kV for the production GTO thyristors, considering the usual factor of three derating for successful commutating. Preconditioning choppers are required to replace the series connection of GTO in a 3 kV DC railway to supply a proper DC link voltage in a two-level inverter [53].

Further development on the structure of inverters gave rise to the application of line connected three-level inverters with a neutral point clamped, which is a feasible alternative in a 3 kV DC railway. In this kind of inverter, 3 voltage states are present in the output terminal, which are 0,  $+V_{DC}/2$  and  $-V_{DC}/2$ . In addition, the virtual inverter switching frequency is four times the actual frequency. The switching frequency thus is typically 100 Hz to 300 Hz and incurs less DC current ripple and a better AC voltage waveform [53].

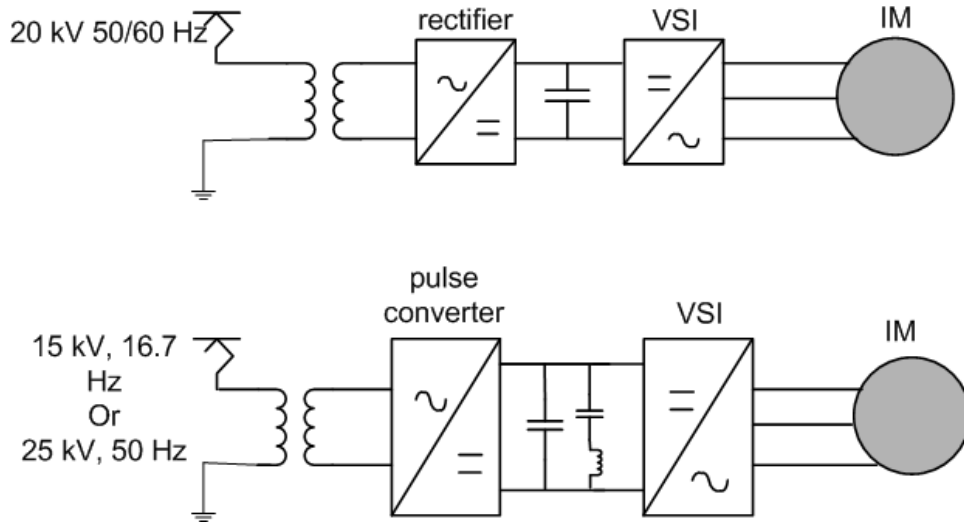


Figure A.7: AC power supply with induction motor drive

### A.2.3 AC-fed Voltage Source Inverter drive

The VSI could be used in AC-fed railways when the AC-DC converter is applied between AC supply lines and VSI, as shown in Figure A.7 [53]. This AC-DC is a four-quadrant pulse converter incorporated with PWM. It inherently has the capability of full regeneration and unity power factor operation.

The condition of the operation can be summarised as follows:

- Power is sent to, or back from the DC link output with unity power factor;
- DC link voltage is maintained constant;
- DC link current demand and line voltage may vary.

The converter has a sinusoidal input voltage in phase with the input current while the output of voltage is required to be constant and ripple free. The AC component of the output current, at twice the supply frequency, is absorbed by a series resonant filter. The circuit diagram of a pulse converter is shown in A.8.

It is proposed that the source voltage  $V_s$  and source current  $I_s$  should be in phase. To align their phases, an inductance is introduced to impose an appropriate phase displacement between source voltage  $V_s$  and the input voltage  $V_{in}$ . If  $V_s$  and

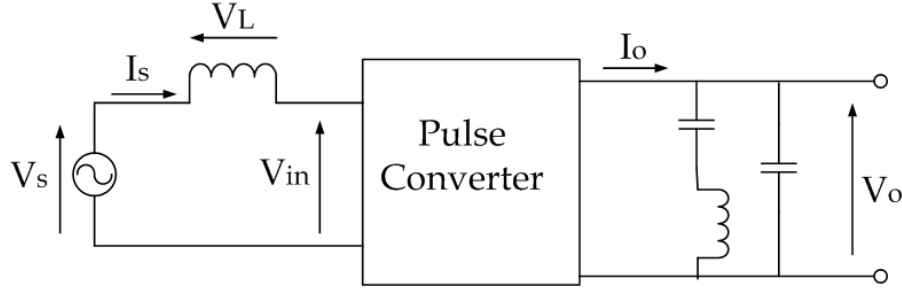


Figure A.8: Circuit diagram of a pulse converter with AC input and DC output.

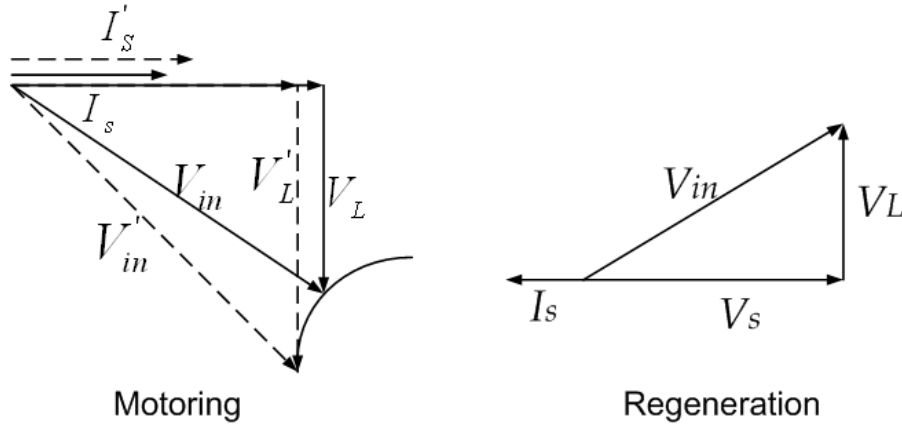


Figure A.9: Phase diagram of AC-DC pulse converter

$I_s$  are known, the amplitude and phase of  $V_{in}$  are also known achieved via PWM converters.

The phase diagram shown in Figure A.9 and Eqn. A.1 demonstrates the relationship between the  $V_s$ ,  $V_{in}$  and  $V_L$  [167].

$$\vec{V}_s + \vec{V}_L = \vec{V}_{in} \quad (A.1)$$

When the  $V_s$  changes, in order to keep the  $I_s$  in phase,  $V_{in}$  should be changed so that the direction of  $V_L$  is shifted in parallel [53]. The  $V_{in}$  is varied by adjusting the modulation ratio of the GTO in relation to the constant link voltage  $V_o$ . The direction of  $I_s$  is reversed from motor mode to regenerative mode.

The application of PWM in the DC-AC converter has several advantages [167]. Firstly, it is easy to achieve the unity power factor by PWM. Secondly, it improves the line current waveform by reducing harmonics. Thirdly, both dynamic and re-

generative braking operations are feasible. Finally, the DC link voltage is larger than the line peak voltage. When the line voltage surges, diodes in the converter could clamp  $V_{in}$  to the DC link voltage, which could in turn protect the devices.

# Appendix B

## Optimisation Techniques

### B.1 Categories of Numerical Optimisation

Depending on the characteristics of the objective function and constraints, numerical optimisation problems in the form of Eqn. 3.1 can be categorised into various sub-fields. Mathematical programming can be categorised into “constrained programming” if certain constraints defined by  $g_i(\cdot)$  and  $h_j(\cdot)$  are met and “unconstrained programming” if  $\mathcal{I} = \emptyset$  and  $\mathcal{E} = \emptyset$ . The nature of the objective function can be regarded as linear demonstrated by the linear programming defined in Def. B.1 or convex demonstrated by the convex programming defined in Def. B.2. As a result, the programming problems can be divided into linear programming and nonlinear programming, convex programming and non convex programming respectively. If the objective function  $f(\cdot)$  is continuous or the variable set only makes sense if it is discrete, the programming problems can then be classified into continuous and discrete programming. Other categorisations are made according to respective emphasis on the characteristics, such as global and local optimisation, stochastic and deterministic optimisation, smooth and non-smooth optimisation etc. [66].

**Definition B.1** (Linear programming). *The optimisation procedure defined in Eqn.*

3.1 is linear programming if the following conditions are satisfied.

$$\begin{aligned}
 f(x) &= c^T x \\
 g_i(x) &= a_i^T x, \quad i = 1, \dots, m \\
 h_j(x) &= a_j^T x \quad j = 1, \dots, l
 \end{aligned} \tag{B.1}$$

The vectors  $c, a_i, a_j \in R^n$  and scalars  $b_i, b_j \in R$  are the parameters which specify the functions  $f(x)$ ,  $g_i(x)$  and  $h_j(x)$ .

Four general assumptions for linear programming problems include proportionality, non-negativity, additivity and linear objective function [69].

**Definition B.2** (Convex optimisation). *The optimisation procedure defined in Eqn. 3.1 is convex optimisation if the following conditions are satisfied.*

$$\begin{aligned}
 f(\alpha x + \beta y) &\leq \alpha f(x_1) + \beta f(x_2) \\
 g_i(\alpha x + \beta y) &\leq \alpha g_i(x_1) + \beta g_i(x_2) \\
 h_j(\alpha x + \beta y) &\leq \alpha h_j(x_1) + \beta h_j(x_2)
 \end{aligned} \tag{B.2}$$

for  $x, y \in R^n$  and  $\alpha, \beta \in R$ , there is  $\alpha + \beta = 1$ ,  $\alpha \geq 0$ ,  $\beta \geq 0$

A classification according to the linearity and convexity of the numerical optimisation problem is shown in Figure B.1. Note that Linear Programming, Second-order Conic Programming (SOCP) and Semidefinite Programming (SDP) are all regarded as Conic Programming.

Numerical algorithms will be categorised into “unconstrained optimisation” algorithms for the cases where  $\mathcal{I} = \emptyset$  and  $\mathcal{E} = \emptyset$  and “constrained optimisation” algorithms otherwise. For more details of linear programming, refer to [69].



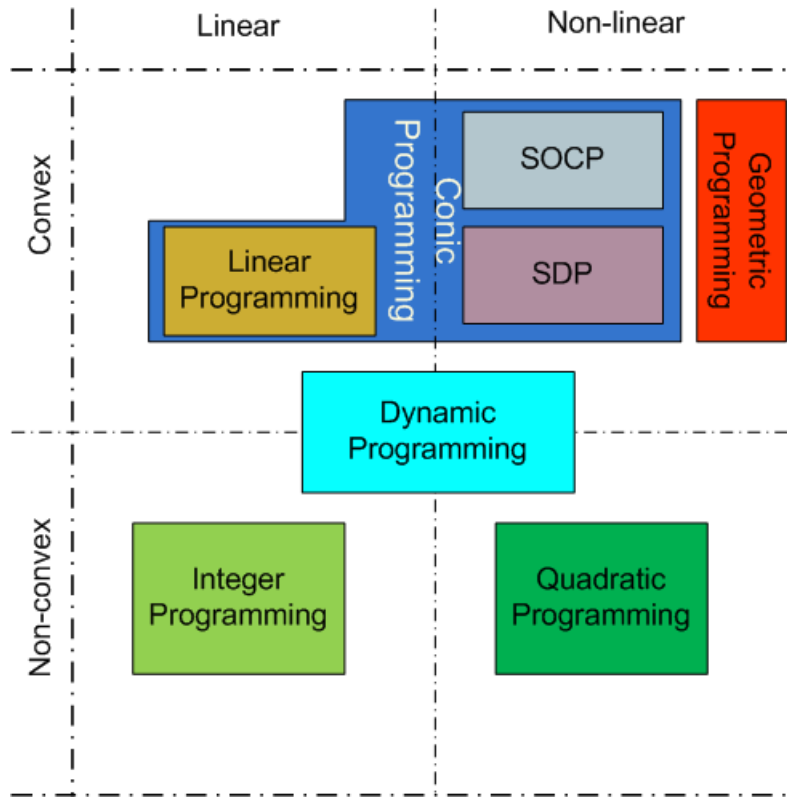


Figure B.1: The classification of the mathematical programming problem

## B.2 Some other metaheuristics

### B.2.1 Simulated annealing

The simulated annealing algorithm was inspired by the “deep and useful connection between statistical mechanics and combinatorial optimisation” [85, 86]. The term “annealing” in metallurgy is a physical process used to heat up and cool down a material in a controlled way in order to increase the effect of crystallisation. When the temperature is high, it is easier for the atoms to become unstuck and to move randomly to other states with higher energy. A controlled, slower cooling process will enable the atoms to gain a higher chance of staying in a state with lower system energy than their initial state.

By analogy, in simulated annealing algorithms, feasible solutions of an optimisation problem correspond to the states of a material. The output of the objective function at a solution corresponds to the system energy at certain state, an optimum

solution corresponds to the minimum energy state.

Given a temperature  $T$ , the probability distribution of system energies  $P(E)$  is proportional to its Boltzmann probability factor:

$$P(E) \propto \exp\left(-\frac{E}{kT}\right) \quad (\text{B.3})$$

Eqn. B.3 indicates that at a lower temperature, lower internal energies  $P(E)$  can still be achieved with a lower possibility. Such statistical probability enables the system to have the potential to escape from a local minimum-energy state [100].

To implement the simulated annealing algorithm for a combinatorial optimisation problem, four components are required [85].

1. The description of a solution of a system;
2. A neighbourhood function which will randomly generate a move or rearrangement of the elements in a solution;
3. An objective function to be optimised;
4. An annealing schedule of temperatures and iteration time length.

Some definitions are made as follows:

- Let  $f(s)$  denote the objective function where  $s$  is one of the solutions and let  $s_{bsf}$  denote the best-so-far solution during optimisation;
- Function INITIALISE-SOLUTION() will return an initial solution;
- Function INITIALISE-PARAMETER() will return a set of initial parameters including: an initial temperature  $T_0$ , termination criterion for iteration at each temperature, e.g. number of iteration  $N_i$ , termination criterion for whole program;
- Function UPDATE-TEMP( $T$ ) will return an updated temperature;

- Function NEIGHBOUR( $s$ ) will return a new solution  $s'$  in the neighbourhood of current solution  $s$ ;
- Function ACCEPT( $s, s', T$ ) will return a final accepted solution for a tentative generated solution  $s'$  at current temperature  $T$ . The pseudo-code of ACCEPT( $s, s'$ ) is shown in Algorithm B.2;
- Function RAND() will return a number  $n_r \in [0, 1]$ .

The pseudo-code for the top level of a typical simulated annealing algorithm is presented in Algorithm B.1.

---

**Algorithm B.1** Pseudo-code for Simulated Annealing

---

```

 $s \leftarrow \text{INITIALISE-SOLUTION}()$ 
INITIALISE-PARAMETER()
 $s_{bsf} \leftarrow s$ 
while Overall termination condition not met do
  while Termination condition at temperature  $T$  not met do
     $s' \leftarrow \text{NEIGHBOUR}(s)$ 
     $s \leftarrow \text{ACCEPT}(s, s', T)$ 
    if  $f(s) < f(s_{bsf})$  then
       $s_{bsf} \leftarrow s$ 
    end if
  end while
   $T \leftarrow \text{UPDATE-TEMP}(T)$ 
end while
return  $s_{bsf}$ 

```

---



---

**Algorithm B.2** Pseudo-code for ACCEPT( $s, s', T$ ) function

---

```

Require:  $p_a \in [0, 1]$ 
 $n_r = \text{RAND}()$  {Randomly generate a number between 0 and 1}
if  $f(s') < f(s)$  then
   $p_a \leftarrow 1$ 
else
   $p_a \leftarrow \exp\left(\frac{f(s) - f(s')}{T}\right)$ 
end if
if  $p_a \geq n_r$  then
  return  $s'$ 
end if
return  $s$ 

```

---

### B.2.2 Tabu Search (TS)

TS is an iterative search algorithm which relies on flexible structure memory [89]. The short-term memory functions of TS constitute a form of “aggressive exploration” in which a “best” move, not necessarily an “improving” move, for an evaluation function is selected subject to certain constraints [87]. Therefore TS makes no reference to a local optimum and is able to search areas beyond a local minimum or a non-feasible area.

There are some policies in TS to prevent any reversal or repetition of moves which can be referred to as the forbidden [87]. A tabu list is used to store solutions already visited in the past “ $T_s$ ” iterations. An appropriate length of “ $T_s$ ”, referred to as the tabu list length, will effectively eliminate the cycling problem whereby the reversal moves occur and progress the search from previous “ $T_s$ ” solutions. Such a tabu list is usually maintained in a “first-in-first-out” style where the oldest solution will be deleted while the newest solution is written. By recording the occurrence of attributes of previous solutions, the tabu restriction strategies can be applied to prevent reversals. The tabu restriction strategies are realised through a comparison of the threshold value and the occurrence frequency in a large number of previous solutions (a frequency-based restriction) or the occurrence rate in a certain number of previous solutions (a recency-based restriction). In other words, the tabu restriction of a solution will be reinforced if the occurrence rate or frequency has passed some threshold.

By contrast, if a solution is high quality and does not cause any recycling problem, it will be accepted as a new solution even if tabu restriction strategies are already met. The strategies adopted here are called “aspiration strategies” which will guide the search in a wider region. Both the  $T_s$  and the aspiration strategies take an important part in freeing strategies in TS as they both relax some restrictions broaden the search space. By promptly deleting the elements in the tabu list, new improved solutions can be re-considered in the future search [100].

While the short-term memory functions are fulfilled by storing the solutions taken in a number of previous iterations, the intermediate term and long-term memory functions operate to intensify and diversify the search by learning from the search behavior in a particular period of the algorithm. The intermediate term memory function records and compares the features of a selected number of best trial solutions in a particular period of search. The common features of all, or a majority, of the best solutions are regarded as regional attributes of good solutions and are used to guide the search for solutions with such attributes in the future. The long-term memory function is applied to diversify the search by guiding it to a region in contrast to the one found thus far during the execution of the algorithm. Different to the random allocation of the starting point, diversification based on the long-term memory tends to learn from the past through an “evaluation criterion”. Based on such an evaluation criterion, the solutions with the features regarded to be popular and effective in the “long term” past will be penalised out of the search in the execution of diversification [87].

Before the algorithmic outline is given in pseudo-code format [84, p.50], some definitions are given as follows.

- Let  $f(s)$  denote the objective function where  $s$  is one of the solutions and let  $s_{bsf}$  denote the best-so-far solution during optimisation;
- Let  $\mathcal{S}$  denote a set of admissible new solutions;
- Function INIT-SOLUTION() returns an initial solution;
- Function INIT-MEMO() will initialise the memory structure of TS;
- Function GENERATE-NEW-SOLUTIONS( $s$ ) returns a new set of accepted solutions based on the tabu list, tabu restriction and tabu aspiration;
- Function BEST-SOLUTION( $s$ )
- Function UPDATE-MEMO() will update the memory structure of TS according to the iterative procedure.

---

**Algorithm B.3** Pseudo-code for Tabu Search

---

```
s ← INIT-SOLUTION()
INIT-MEMO()
 $s_{bsf} \leftarrow s$ 
while Termination conditions not met do
   $\mathcal{S} \leftarrow \text{GENERATE-NEW-SOLUTIONS}(s)$ 
   $s \leftarrow \text{BEST-SOLUTION}(\mathcal{S})$ 
  UPDATE-MEMO()
  if  $f(s) < f(s_{bsf})$  then
     $s_{bsf} \leftarrow s$ 
  end if
end while
return  $s_{bsf}$ 
```

---

# Appendix C

## Publications

Publications generated to date based on the research results in this thesis are listed as follows.

1. S. Lu, S. Hillmansen, and C. Roberts, "*A Power-Management Strategy for Multiple-Unit Railroad Vehicles*," Vehicular Technology, IEEE Transactions on, vol. 60, pp. 406-420, 2011
2. S. Lu, S. Hillmansen, and C. Roberts, "*Power management strategy study for a multiple unit train*," IET Seminar Digests, vol. 2010, pp. 29-29, 2010.
3. S. Lu, D. H. Meegahawatte, S. Guo, S. Hillmansen, C. Roberts, and C.J. Goodman, "*Analysis of Energy Storage Devices in Hybrid Railway Vehicles*," in International conference on railway engineering, IET 2008 Hongkong, 2008.

## References

- [1] Z. Utlu and A. Hepbasli, “Assessment of the energy utilization efficiency in the Turkish transportation sector between 2000 and 2020 using energy and energy analysis method,” *Energy Policy*, vol. 34, pp. 1611–1618, Sep. 2006.
- [2] R. C. Hibbeler, *Mechanics of materials*. Singapore; London: Prentice Hall, 7th ed., 2008.
- [3] F. Browand, R. McCallen, J. Ross, A. Orellano, and S. Sperling, *The Aerodynamics of Heavy Vehicles II: Trucks, Buses, and Trains*, vol. 41 of *Lecture Notes in Applied and Computational Mechanics*, ch. Aerodynamic Improvements and Associated Energy Demand Reduction of Trains, pp. 219–231. Springer Berlin / Heidelberg, 2009.
- [4] “Delivering a sustainable railway - white paper cm 7176.” Government white paper, 2007.
- [5] M. Ogasu, “Energy saving and environmental measures in railway technologies: example with hybrid electric railway vehicles,” *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 3, pp. 15–20, Jan. 2008.
- [6] L. Kantorovich, *Mathematical methods in the organization and planning of production*. Publication House of the Leningrad State University, 1939.
- [7] B. P. Rochard and F. Schmid, “Benefits of lower-mass trains for high speed rail operations,” *Proceedings of the Institution of Civil Engineers: Transport*, vol. 157, no. 1, pp. 51–64, 2004.



---

## REFERENCES

---

- [8] P. C. Tan, P. C. Loh, and D. G. Holmes, "Optimal impedance termination of 25-kv electrified railway systems for improved power quality," *IEEE Transactions on Power Delivery*, vol. 20, pp. 1703 – 1710, Apr. 2005.
- [9] Z. W. Zhang, B. Wu, J. S. Kang, and L. F. Luo, "A multi-purpose balanced transformer for railway traction applications," *IEEE Transactions on Power Delivery*, vol. 24, pp. 711–718, Apr. 2009.
- [10] C. B. Park, B. S. Lee, H. W. Lee, S. Y. Kwon, and H. J. Park, "Air-gap control system of a linear induction motor for a railway transit," in *International Conference on Electrical Machines, vols I- 4*, pp. 2000–2003, 2008.
- [11] C. H. Bae, M. S. Han, Y. K. Kim, C. Y. Choi, and S. J. Jang, "Simulation study of regenerative inverter for DC traction substation," in *Proceedings of the Eighth International Conference on Electrical Machines and Systems*, vol. 2, pp. 1452–1456, 2005.
- [12] B. Mellitt, Z. S. Mouneimne, and C. J. Goodman, "Simulation study of DC transit systems with inverting substations," *Electric Power Applications, IEE Proceedings B*, vol. 131, no. 2, pp. 38–50, 1984.
- [13] A. Adinolfi, R. Lamedica, C. Modesto, A. Prudenzi, and S. Vimercati, "Experimental assessment of energy saving due to trains regenerative braking in an electrified subway line," *IEEE Transactions on Power Delivery*, vol. 13, pp. 1536–1542, Oct. 1998.
- [14] A. Adinolfi, R. Lamedica, C. Modesto, A. Prudenzi, and S. Vimercati, "Experimental assessment of energy saving due to trains regenerative braking in an electrified subway line," in *Industrial and Commercial Power Systems Technical Conference*, pp. 211–216, IEEE, 1997.
- [15] E. Agenjos, A. Gabaldon, F. Franco, R. Molina, S. Valero, M. Ortiz, and R. Gabaldon, "Energy efficiency in railways: energy storage and electric

---

## REFERENCES

---

- generation in diesel electric locomotives,” *IET Conference Publications*, no. CP550, pp. 402–402, 2009.
- [16] D. Iannuzzi, “Improvement of the energy recovery of traction electrical drives using supercapacitors,” in *13th Power Electronics and Motion Control Conference*, pp. 1469 –1474, Sept. 2008.
- [17] P. A. Flaherty, “Multi-stage hybrid drives for traction applications,” in *Proceedings of the ASME/IEEE Joint Rail Conference*, pp. 171 – 175, Mar. 2005.
- [18] S. Oettich, T. Albrecht, and S. Scholz, “Improvements of energy efficiency of urban rapid rail systems,” *Urban Transport X*, vol. 16, pp. 573–582, 2004.
- [19] R. Liu and L. M. Golovitcher, “Energy-efficient operation of rail vehicles,” *Transportation Research Part A: Policy and Practice*, vol. 37, no. 10, pp. 917–932, 2003.
- [20] T. K. Ho and T. H. Yeung, “Railway junction traffic control by heuristic methods,” *IEE Proceedings- Electric Power Applications*, vol. 148, no. 1, pp. 77–84, 2001.
- [21] Y.-H. Chou, *Distributed approach for rescheduling railway traffic in the event of disturbances*. PhD, School of Electronic, Electrical and Computer Engineering, The University of Birmingham, 2009.
- [22] G. J. Hull, “Simulation of energy efficiency improvements on commuter railways,” Master of Philosophy, School of Electronic, Electrical and Computer Engineering, the University of Birmingham, Dec. 2009.
- [23] H.-S. Hwang, “Control strategy for optimal compromise between trip time and energy consumption in a high-speed railway,” *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 28, pp. 791–802, Nov. 1998.

---

## REFERENCES

---

- [24] K. K. Wong and T. K. Ho, "Coast control for mass rapid transit railways with searching methods.," *IEE Proceedings -Electric Power Applications*, vol. 151, no. 3, pp. 365–376, 2004.
- [25] C. Chang and S. Sim, "Optimising train movements through coast control using genetic algorithms," *IEE Proceeding -Electric Power Applications*, vol. 144, no. 1, pp. 65–73, 1997.
- [26] Y. V. Bocharnikov, A. M. Tobias, C. Roberts, S. Hillmanssen, and C. J. Goodman, "Optimal driving strategy for traction energy saving on DC suburban railways," *Electric Power Applications, IET*, vol. 1, no. 5, pp. 675–682, 2007.
- [27] H.-J. Chuang, C.-S. Chen, C.-H. Lin, C.-H. Hsieh, and C.-Y. Ho, "Design of optimal coasting speed for saving social cost in mass rapid transit systems," in *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, pp. 2833–2839, 2008.
- [28] T. Albrecht, *Railway timetable & traffic: analysis, modelling and simulation*, ch. Energy-Efficient Train Operation, pp. 83–105. Eurailpress | DVV Rail Media (DVV Media Group GmbH), 2008.
- [29] P. Howlett, "The optimal control of a train," *Annals of Operation Research*, vol. 98, pp. 65–87, 2000.
- [30] E. Khmelnitsky, "On an optimal control problem of train operation," *IEEE transactions on automatic control*, vol. 45, no. 7, pp. 1257–1266, 2000.
- [31] H. Ko, T. Koseki, and M. Miyatake, "Application of dynamic programming to the optimization of the running profile of a train," in *Computers in Railway IX* (J. Allan, C. Brebbia, R. Hill, G. Sciutto, and S. Sone, eds.), vol. 15, pp. 103–112, The Wessex Institute, 2004.

---

## REFERENCES

---

- [32] S. Effati and H. Roohparvar, "The minimization of the fuel costs in the train transportation," *Applied Mathematics and Computation*, vol. 175, pp. 1415–1431, Apr. 2006.
- [33] W. Liu, Q. Li, and B. Tang, "Energy saving train control for urban railway train with multi-population genetic algorithm," in *International forum on information technology and applications*, pp. 58–63, IEEE, IEEE computer society, 2009.
- [34] T. Albrecht, "Reducing power peaks and energy consumption in rail transit systems by simultaneous train running time control," *Computers in Railway Six*, vol. 15, pp. 885–894, 2006.
- [35] R. Takagi and S. Sone, "More intelligent DC railway electrical-power systems with traction power-control," *Electrical Engineering in Japan*, vol. 114, pp. 91–101, Dec. 1994.
- [36] B. Sansó and P. Girard, "Instantaneous power peak reduction and train scheduling desynchronization in subway systems," *Transportation Science*, vol. 31, no. 4, p. 312, 1997.
- [37] S. P. Gordon and D. G. Lehrer, "Coordinated train control and energy management control strategies," in *Proceedings of the 1998 ASME/IEEE Joint Railroad Conference* (D. Stone and D. Haluza, eds.), pp. 165–176, AMER SOC Mechanical Engineers, 1998.
- [38] B. Mellitt, C. Goodman, and R. Arthurton, "Simulation studies of energy saving with chopper control on the Jubilee line," *Proceeding of IEE*, vol. 125, no. 4, pp. 304–310, 1978.
- [39] S. Acikbas and M. T. Soylemez, "Parameters affecting braking energy recuperation rate in dc rail transit," in *ASME/IEEE Joint Rail Conference/ASME Internal Combustion Engine Division Spring Technical Con-*

---

## REFERENCES

---

- ference*, pp. 263–268, ASME, Transportat Div; ASME, Internal Combustion Engine Div, AMER SOC Mechanical Engineers, Three Park Avenue, New York, NY 10016-5990 USA, 2007.
- [40] B. Powell, K. Bailey, and S. Cikanek, “Dynamic modeling and control of hybrid electric vehicle powertrain systems,” *Control System Magazine, IEEE*, vol. 18, no. 5, pp. 17–33, 1998.
- [41] A. R. Miller, J. Peters, B. E. Smith, and O. A. Velev, “Analysis of fuel cell hybrid locomotives,” *Journal of Power Sources*, vol. 157, no. 2, pp. 855–861, 2006.
- [42] A. R. Miller, K. S. Hess, D. L. Barnes, and T. L. Erickson, “System design of a large fuel cell hybrid locomotive,” *Journal of Power Sources*, vol. 173, pp. 935–942, Nov. 2007.
- [43] T. Horiba, K. Hironaka, T. Matsumura, T. Kai, M. Koseki, and Y. Muranaka, “Manganese-based lithium batteries for hybrid electric vehicle applications,” *Journal of Power Sources*, vol. 119–121, pp. 893–896, 2003.
- [44] C. Deng, P.-F. Shi, and S. Zhang, “Development of novel bipolar nickel/metal hybrid batteries for hybrid electric vehicles,” *Chinese Journal of Chemistry*, vol. 23, no. 2, pp. 221–224, 2005.
- [45] D. Ohms, M. Kohlhase, G. Benczur-Urmossy, K. Wiesener, and J. Harmel, “High performance nickel-metal hybrid battery in bipolar stack design,” *Journal of Power Sources*, vol. 105, no. 2, pp. 120–126, 2002.
- [46] Anonymous, *Automotive Electronics*. John Wiley Sons Ltd, 5th ed., 2007.
- [47] Y. S. W. K. T. Chau, “Overview of power management in hybrid electric vehicles,” *Energy Conversion and Management*, vol. 43, pp. 1953–1968, October 2002.

---

## REFERENCES

---

- [48] S. Lu, S. Hillmansen, and C. Roberts, “Power management strategy study for a multiple unit train,” *IET Seminar Digests*, vol. 2010, no. 13342, pp. 29–29, 2010.
- [49] S. Lu, S. Hillmansen, and C. Roberts, “A power management strategy for multiple unit railroad vehicles,” *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 406–420, 2011.
- [50] N. Donovan, “Transpennine trains get greener,” *Railway Gazette International*, vol. 164, pp. 715–715, Sep. 2008.
- [51] L. Guzzella and A. Sciarretta, *Vehicle Propulsion Systems, Introduction to Modeling and Optimization*. Springer Berlin Heidelberg, second edition ed., 2007.
- [52] S. Hillmansen, “Sustainable traction drives,” in *IET Seminar Digest*, pp. 255–265, Institution of Engineering and Technology, 2009.
- [53] R. J. Hill, “Electric railway traction: Part 2 traction drives with three-phase induction motors,” *Power Engineering Journal*, vol. 8, pp. 143–152, Jun. 1994.
- [54] R. J. Hill, “Electric railway traction: Part 1 electric traction and DC traction motor drives,” *Power Engineering Journal*, vol. 8, pp. 47–56, Feb. 1994.
- [55] H. I. Andrews, *Railway traction : the principles of mechanical and electrical railway traction*. Studies in mechanical engineering 5, Amsterdam ; Oxford : Elsevier Science, 1986.
- [56] R. Hill, “Electric railway traction. part 3. traction power supplies,” *Power Engineering Journal*, vol. 8, pp. 275 –286, Dec. 1994.
- [57] Anonymous, “Britain’s transport infrastructure: Rail electrification,” tech. rep., Department for Transportation, 2009. [www.dft.org.uk](http://www.dft.org.uk).

---

## REFERENCES

---

- [58] J. Bruce, “The diesel engine in railway transportation,” *The New Zealand Railway Magazine*, vol. 1, pp. 14–16, Aug. 1926.
- [59] T. Ogawa, “Design estimation of the hybrid power source railway vehicle based on the multiple optimization by the dynamic programming,” *Electrical and Electronic Engineering, IEEJ Transactions on*, no. 3, pp. 48–55, 2008.
- [60] C. Jefferson and J. Marquez, “Hybrid powertrain for a light rail vehicle,” in *39th International Universities Power Engineering Conference* (J. Marquez, ed.), vol. 3, pp. 1267–1273, 2004.
- [61] J. S. Chen and M. Salman, “Learning energy management strategy for hybrid electric vehicles,” in *IEEE Conference on Vehicle Power and Propulsion* (M. Salman, ed.), pp. 68–73, 2005.
- [62] S. Sone, “Wayside and on-board storage can capture more regenerated energy,” *Railway Gazette*, 2007. <http://www.railwaygazette.com>.
- [63] D. Meegahawatte, S. Hillmanssen, C. Roberts, M. Falco, A. McGordon, and P. Jennings, “Analysis of a fuel cell hybrid commuter railway vehicle,” *Journal of Power Sources*, vol. 195, no. 23, pp. 7829–7837, 2010.
- [64] S. J. Chapman, *Electric Machinery Fundamentals*. Australia: The McGraw-Hill companies, 4 ed., 2005. Electric Motors.
- [65] R. Gabriel, W. Leonhard, and C. J. Nordby, “Field-oriented control of a standard ac motor using microprocessors,” *IEEE Transactions on Industry Applications*, vol. IA-16, pp. 186 –192, Mar. 1980.
- [66] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering, Springer New York, 2nd ed., 2006.

---

## REFERENCES

---

- [67] M. W. Jeter, *Mathematical Programming: An Introduction to Optimization*. Monographs and textbooks in pure and applied mathematics, New York : M. Dekker, Inc., 1986.
- [68] A. A. Hopgood, *Intelligent Systems for Engineers and Scientists, Second Edition*. CRC Press, 2009.
- [69] G. B. Dantzig, *Linear Programming and Extensions*. Princeton University Press, 1963.
- [70] G. B. Arfken, *Mathematical methods for physicists*. London: Academic Press, 1985. QA 401/A.
- [71] S. Dreyfus, “Recharh Bellman on the birth of dynamic programming,” *Operational Research*, vol. 50, no. 1, pp. 48–51, 2002.
- [72] A. Lew and H. Mauch, *Dynamic programming: a computational tool*. Studies in computational intelligence, Springer, 2007.
- [73] R. E. Bellman, *Dynamic programming*. Rand Corporation research study, Princeton University Press, 1957.
- [74] R. Bellman, “A Markovian decision process,” *Journal of mathematics and mechanics*, vol. 6, 1957.
- [75] R. A. Howard, *Dynamic programming and Markov processes*. The M.I.T. Press, 1960.
- [76] B. Givan and R. Parr, “Introduction to Markov decision processes.” [www.cs.rice.edu/~vardi/dag01/givan1.pdf](http://www.cs.rice.edu/~vardi/dag01/givan1.pdf). Achieved on 11/08/2010.
- [77] T. H. Corman, R. L. Leiserson, Charles E. and Rivest, and C. Stein, *Introduction to Algorithms*. The MIT Press, 2nd ed., Sep. 2001.
- [78] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.



---

## REFERENCES

---

- [79] D. B. Wagner, “Dynamic programming,” *The Mathematica Journal*, vol. 5, no. 4, pp. 42–51, 2006.
- [80] S. R. Eddy, “What is dynamic programming,” *Nature Biotechnology*, vol. 22, no. 7, pp. 909–910, 2004.
- [81] F. J. Bonnans, C. J. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical optimization. Theoretical and Practical Aspects*, Springer Berlin Heidelberg, 2nd ed., 2006.
- [82] F. Glover, “Future paths for integer programming and links to artificial intelligence,” *Computers and Operations Research*, vol. 13, pp. 533–549, 1986.
- [83] F. Glover and G. Kochenberger, eds., *Handbook of Metaheuristics*, vol. 57. Kluwer Academic Publishers, 2002.
- [84] M. Dorigo and S. Thomas, *Ant colony optimisation*. The MIT Press, 2004.
- [85] S. Kirkpatrick, C. Gelatt(Jr.), and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, pp. 671–680, May 1983.
- [86] V. Cerný, “A thermodynamical approach to the traveling salesman problem,” *Journal of Optimization Theory and Applications*, vol. 45, no. 1, pp. 41–45, 1985.
- [87] F. Glover, “Tabu search-part I,” *ORSA Journal on Computing*, vol. 1, no. 3, pp. 190–206, 1989.
- [88] F. Glover, “Tabu search-part II,” *ORSA Journal on Computing*, vol. 2, no. 1, pp. 4–32, 1990.
- [89] F. Glover, “Tabu search: a tutorial,” *Interfaces*, vol. 20, no. 4, pp. 74–94, 1990.

---

## REFERENCES

---

- [90] C. Voudouris, *Guided local search for combinatorial optimization problems*. PhD, Department of Computer Science, University of Essex, Colchester, UK, 1997.
- [91] C. Voudouris and E. Tsang, “Guided local search and its application to the traveling salesman problem,” *European Journal of Operational Research*, vol. 113, pp. 469–499, Mar. 1999.
- [92] T. A. Feo and M. G. C. Resende, “A probabilistic heuristic for a computationally difficult set covering problem,” *Operational research letters*, vol. 8, pp. 67–71, 1989.
- [93] T. A. Feo and M. G. C. Resende, “Greedy randomized adaptive search procedures,” *Journal of Global Optimization*, vol. 6, pp. 109–134, 1995.
- [94] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, 1st ed., 1992.
- [95] H. R. Lourenço, O. C. Martin, and T. Stützle, “Iterated local search,” Economics Working Papers 513, Department of Economics and Business, Universitat Pompeu Fabra, 2000.
- [96] F. Glover, “Heuristics for integer programming using surrogate constraints,” *Decision Sciences*, vol. 8, no. 1, pp. 156–166, 1977.
- [97] T. C. Jones, *Evolutionary Algorithms, Fitness Landscapes and Search*. PhD, University of New Mexico, 1995.
- [98] X.-S. Yang, *Biology-Derived Algorithms in Engineering Optimization*, ch. 32, pp. 1–12. Taylor & Francis Group, LLC, 2006.
- [99] E. Alba and C. Cotta, *Handbook of bioinspired algorithms and applications*, ch. 1, pp. 1–19. Taylor & Francis Group, LLC, 2006.

---

## REFERENCES

---

- [100] D. T. Pham and D. Karaboga, *Intelligent optimisation techniques*. Springer, 2000.
- [101] D. Whitley, “A genetic algorithm tutorial,” *Statistics and Computing*, vol. 4, pp. 65–85, 1994.
- [102] G. Beni and J. Wang, “Swarm intelligence in cellular robotic systems,” in *Proceedings of NATO Advanced Workshop on Robots and Biological Systems*, vol. 102, 1989.
- [103] M. Zlochin, M. Birattari, N. Meuleau, and M. Dorigo, “Model-based search for combinatorial optimisation: A critical survey,” *Annals of Operations Research*, vol. 131, pp. 373–395, 2004.
- [104] M. Dorigo, *Optimization, Learning and Natural Algorithms*. PhD, Politecnico di Milano, 1992.
- [105] V. Maniezzo, L. M. Gambardella, and F. D. Luigi, *New Optimization Techniques in Engineering*, ch. Ant Colony Optimization, pp. 101–117. Springer-Verlag Berlin Heidelberg, 2004.
- [106] A. Coloni, M. Dorigo, and V. Maniezzo, “Distributed optimization by ant colonies,” in *Proceedings of European Conference on Artificial Life*, pp. 134–142, European Conference on Artificial Life, Elsevier publishing, 1991.
- [107] M. Dorigo, V. Maniezzo, and A. Coloni, “Positive feedback as a search strategy,” Technical Report 91-016, Politecnico di Milano, 1991.
- [108] M. Dorigo, V. Maniezzo, and A. Coloni, “The ant system: Optimization by a colony of cooperating agents,” *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, vol. 26, pp. 29–41, 1996.
- [109] L. M. Gambardella and M. Dorigo, “Ant-q: A reinforcement learning approach to the traveling salesman problem,” in *Proceedings of the Twelfth*

---

## REFERENCES

---

- International Conference on Machine Learning*, pp. 252–260, Morgan Kaufmann, 1995.
- [110] M. Dorigo and L. M. Gambardella, “Ant colony system: A cooperative learning approach to the traveling salesman problem,” *IEEE Transactions on Evolutionary Computation*, vol. 1, pp. 53–66, 1996.
- [111] T. Stützle and H. Hoos, “Improving the ant system: A detailed report on the max-min ant system,” tech. rep., TU Darmstadt, Germany, 1996.
- [112] T. Stutzle and H. H. Hoos, “Max-min ant system,” *Future Generation Computer Systems*, vol. 16, no. 9, pp. 889–914, 2000.
- [113] B. Bullnheimer, R. F. Hartl, and C. Strauß, “A new rank based version of the ant system - a computational study,” *Central European Journal for Operations Research and Economics*, vol. 7, pp. 25–38, 1997.
- [114] V. Maniezzo, “Exact and approximate nondeterministic tree-search procedures for the quadratic assignment problem,” *INFORMS Journal of Computing*, vol. 11, no. 4, pp. 358–369, 1999.
- [115] C. Blum and M. Dorigo, “The hyper-cube framework for ant colony optimization,” *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 34, pp. 1161–1172, Apr. 2004.
- [116] C. J. Goodman, L. K. Siu, and T. K. Ho, “A review of simulation models for railway systems,” in *International conference on Developments in Mass Transit Systems*, pp. 80–85, IEE, Apr. 1998. Train simulation study.
- [117] C. J. Goodman, “Basic physics of traction and train performance calculations.” Courses note, MSc programme in railway systems engineering and integration, University of Sheffield, Oct. 1995.

---

## REFERENCES

---

- [118] B. Mao, W. Jia, S. Chen, and J. Liu, "A computer-aided multi-train simulator for rail traffic," in *IEEE International Conference on Vehicular Electronics and Safety*, pp. 1 –5, 13-15 2007.
- [119] D. Yu, K. Lo, X. D. Wang, C. G. Yin, and D. L. Huang, "Analysis of dynamic mrts traction power supply system based on dependent train movement simulation," in *Proceedings of the ASME/IEEE Joint Rail Conference*, pp. 153–161, 6-8 2004.
- [120] P. Martin, "Train performance and simulation," in *Simulation Conference Proceedings*, vol. 2, pp. 1287 –1294, 1999.
- [121] S. Hillmansen and C. Roberts, "Energy storage devices in hybrid railway vehicles: A kinematic analysis," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 221, no. 1, pp. 135–143, 2007.
- [122] C. S. Chang, C. S. Chua, H. B. Quek, X. Y. Xu, and S. L. Ho, "Development of train movement simulator for analysis and optimisation of railway signalling systems," in *International Conference on Developments in Mass Transit Systems*, pp. 243–248, 20-23 1998.
- [123] W. Jia, S. Chen, T. Ho, B. Mao, and Y. Bai, "A heuristic algorithm for fixed train runtime," in *International Conference on Railway Engineering - Challenges for Railway Transportation in Information Age*, pp. 1 –7, 25-28 2008.
- [124] K. K. Wong and T. k. Ho, "Dwell-time and run-time control for DC mass rapid transit railways," *Electric Power Applications, IET*, vol. 1, no. 6, pp. 956–966, 2007.
- [125] W. H. Zhang, J. Z. Chen, X. J. Wu, and X. S. Jin, "Wheel/rail adhesion and analysis by using full scale roller rig," *Wear*, vol. 253, no. 1-2, pp. 82–88, 2002.

---

## REFERENCES

---

- [126] J. Tunley, "Sand on the line - a recent history," in *IMeche Conference Transactions*, Railway Safety, Nov. 2001.
- [127] C. J. Goodman, "Basic physics of traction and train performance calculations." MSc programme in railway systems engineering and integration courses note - the University of Sheffield, 1995.
- [128] A. Baron, M. Mossi, and S. Sibilla, "The alleviation of the aerodynamic drag and wave effects of high-speed trains in very long tunnels," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 89, pp. 365–401, April 2001.
- [129] B. P. Rochard and F. Schmid, "A review of methods to measure and calculate train resistances," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 214, no. 4, pp. 185–199, 2000.
- [130] C. J. Goodman, "Modelling and simulation." Courses note, MSc programme in railway systems engineering and integration, University of Birmingham.
- [131] I. Mitchell, "The sustainable railway use of advisory systems for energy savings," *News view*, pp. 2–8, Dec. 2009. Technical Paper.
- [132] S.-H. Han, S.-G. Lee, S.-G. Kim, and W.-D. Lee, "Design of optimal control for automatic train operation system in emu," in *International Conference on Control, Automation and Systems*, pp. 394–397, IEE, 2001.
- [133] K. K. Wong and T. K. Ho, "Dynamic coast control of train movement with genetic algorithm," *International Journal of Systems Science*, vol. 35, pp. 835–846, Oct. 2004.
- [134] L. M. Hocking, *Optimal Control: An Introduction to the Theory with Applications*. Oxford Applied Mathematics and Computing Science Series, Clarendon Press. Oxford, 1991.

---

## REFERENCES

---

- [135] M. Athans and P. L. Falb, *Optimal control: An Introduction to the Theory and Its Applications*. McGraw-Hill Book Company, 1966.
- [136] C. Johnson and J. Gibson, "Singular solutions in problems of optimal control," *Automatic Control, IEEE Transactions on*, vol. 8, pp. 4 – 15, Jan. 1963.
- [137] L. M. Sonneborn and F. S. V. Vleck, "The bang-bang principle for linear control systems," *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, vol. 2, no. 2, pp. 151–159, 1964.
- [138] Mathworks, "Genetic algorithm options." <http://www.mathworks.com/help/toolbox/gads/f6174dfi10.html>. Accessed on 12th Sep. 2010.
- [139] Y. Cai, M. R. Irving, and S. H. Case, "Iterative techniques for the solution of complex DC-rail-traction systems including regenerative braking," *IEE Proceedings-Generation Transmission and Distribution*, vol. 142, no. 5, pp. 445–452, 1995.
- [140] C. J. Goodman and L. K. Siu, "DC railway network solutions by diakoptics," in *Proceedings of ASME/IEEE Joint Railroad Conference*, pp. 103–110, ASME/IEEE, 1994.
- [141] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*. Computer science and applied mathematics, Academic press, Inc., 1970.
- [142] C. L. Pires, S. I. Nabeta, and J. R. Cardoso, "Iccg method applied to solve dc traction load flow including earthing models," *IET Electric Power Applications*, vol. 1, pp. 193–198, Mar. 2007.
- [143] Y. Ding, F.-m. Zhou, Y. Bai, T.-k. Ho, and Y.-f. Fung, "Parallel computing for multi-train movement simulation on electrified railway," in *Second International Conference on Information and Computing Science*, vol. 4, pp. 280–283, 21-22 2009.

---

## REFERENCES

---

- [144] M. Chymera, A. C. Renfrew, and M. Barnes, "Analysis of power quality in a dc tram system," in *The 3rd IET International Conference on Power Electronics, Machines and Drives, 2006*, IET, 2006.
- [145] J.O.Bird, *Electrical circuit theory and technology*. Elsevier Science, 2nd edition ed., 2003.
- [146] S. Lu, D. Meegahawatte, S. Guo, S. Hillmanssen, C. Roberts, and C. Goodman, "Analysis of energy storage devices in hybrid railway vehicles," in *International conference on railway engineering, IET*, The Institution of Engineering of Technology, 2008.
- [147] J. M. Miller, "Power electronics in hybrid electric vehicle applications," in *Conference Proceedings - IEEE Applied Power Electronics Conference and Exposition - APEC*, (J-N-J Miller Design Services, P.L.C., 3573 E. Gatzke Road, Cedar, MI 49621, United States), pp. 23–29, IEEE, 2003.
- [148] A. D. Napoli, F. Crescimbeni, L. Solero, F. Caricchi, and G. F. Capponi, "Multiple-input DC-DC power converter for power-flow management in hybrid vehicles," in *Conference Proceedings - IEEE Industry Applications Conference 37th IAS Annual Meeting*, pp. 1578–1585, IEEE, 2002.
- [149] F. R. Salmasi, "Control strategies for hybrid electric vehicles: Evolution, classification, comparison, and future trends," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 2390–2404, 2007. 0018-9545.
- [150] P. Caratozzolo, M. Serra, and J. Riera, "Energy management strategies for hybrid electric vehicles," in *IEEE International Electric Machines and Drives Conference* (J.Riera, ed.), pp. 241–248, IEEE, IEEE International Electric Machines and Drives Conference, 2003.



---

## REFERENCES

---

- [151] T. Hofman, M. Steinbuch, R. V. Druten, and A. Serrarens, “Rule-based energy management strategies for hybrid vehicles,” *International Journal of Electric and Hybrid Vehicles*, vol. 1, no. 1, pp. 71–94, 2007.
- [152] N. A. Kheir, M. A. Salman, and N. J. Schouten, “Emissions and fuel economy trade-off for hybrid vehicles using fuzzy logic,” *Mathematics and Computers in Simulation*, vol. 66, no. 2-3, pp. 155–172, 2004.
- [153] B. M. Baumann, G. Washington, B. C. Glenn, and G. Rizzoni, “Mechatronic design and control of hybrid electric vehicles,” *IEEE/ASME Transactions on Mechatronics*, vol. 5, pp. 58–72, Mar. 2000.
- [154] H.-D. Lee and S. Seung-Ki, “Fuzzy-logic-based torque control strategy for parallel-type hybrid electric vehicle,” *IEEE Transaction on Industrial Electronics*, vol. 45, no. 4, pp. 625–632, 1998.
- [155] S. Delprat, T. Guerra, and J. Rimaux, “Optimal control of a parallel powertrain: from global optimization to real time control strategy,” in *IEEE 55th Vehicular Technology Conference*, vol. 4, pp. 2082–2088, Fall 2002.
- [156] G. Paganelli, S. Delprat, T. Guerra, J. Rimaux, and J. Santin, “Equivalent consumption minimization strategy for parallel hybrid powertrains,” in *IEEE 55th Vehicular Technology Conference*, vol. 4, pp. 2076–2081, IEEE, Spring 2002.
- [157] T. Ogawa, H. Yoshihara, S. Wakao, K. Kondo, and M. Kondo, “Energy consumption analysis of fc-edlc hybrid railway vehicle by dynamic programming,” in *2007 European Conference on Power Electronics and Applications*, 2007.
- [158] E. D. Tate and S. P. Boyd, “Finding ultimate limits of performance for hybrid electric vehicles,” in *SAE Future Transportation Technology Conference and*

---

## REFERENCES

---

- Exposition*, vol. 109, SAE International, Warrendale, Pennsylvania, USA, Aug. 2000.
- [159] B. Huang, X. Shi, and Y. Xu, “Parameter optimization of power control strategy for series hybrid electric vehicle,” in *IEEE Congress on Evolutionary Computation*, pp. 1989–1994, IEEE, 2006.
- [160] M.-G. Morteza and A. Poursamad, “Application of genetic algorithm for simultaneous optimisation of HEV component sizing and control strategy,” *International Journal of Alternative Propulsion*, vol. 1, no. 1, pp. 63–78, 2006.
- [161] A. Sciarretta, M. Back, and L. Guzzella, “Optimal control of parallel hybrid electric vehicles,” *IEEE Transactions on Control Systems Technology*, vol. 12, no. 3, pp. 352–363, 2004.
- [162] S. Delprat, T. M. Guerra, and J. Rimaux, “Control strategies for hybrid vehicles: optimal control,” in *IEEE 56th Vehicular Technology Conference*, vol. 3, pp. 1681–1685, IEEE, Fall 2002.
- [163] Anonymous, “Disiro UK DMU Class 185,” technique report, Siemens Transportation, 2004. [www.siemenstransportation.co.uk/pdfs/185.pdf](http://www.siemenstransportation.co.uk/pdfs/185.pdf).
- [164] J. Klapp, J. L. Cervantes-Cota, and J. F. C. Alcalá, eds., *Towards a Cleaner Planet*. Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, 2007.
- [165] Y. Zhu, Y. Chen, G. Tian, H. Wu, and Q. Chen, “A four-step method to design an energy management strategy for hybrid vehicles,” in *2004 American Control Conference*, Jun. 2004.
- [166] R. H. Byrd, M. E. Hribar, and J. Nocedal, “An interior point algorithm for large scale nonlinear programming,” *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1997.

## REFERENCES

---

- [167] C. W. Lander, *Power Electronics*. McGraw-Hill Publishing Companies, 3 ed., 1993.