MULTIVARIABLE PREDICTIVE CONTROL OF A TMP PLANT

By

Huaijing Du

B.A. Sc. Beijing University of Chemical Technology (1983)
M.A. Sc. Beijing University of Chemical Technology (1988)

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Doctor of Philosophy

in
THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
October 1998
© Huaijing Du, 1998
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical and Computer Engineering

The University of British Columbia
Vancouver, Canada

Date 1998. 10. 22
Abstract

This thesis describes the development of a novel control strategy for a two-stage thermo-mechanical pulping (TMP) plant. Desired pulp quality is achieved by selecting the set-points of specific energy and refining intensity at each stage. The targets of specific energy and refining intensity are obtained through the control of motor load, production rate and refining consistency by manipulating closing pressure, chip flow rate and dilution flow rate at the inlet of each stage. A constrained predictive controller is developed based on the generalized predictive control (GPC) algorithm because of its simplicity, ease of use and ability to handle problems in one algorithm. Future control actions are determined by minimizing the predicted errors without violating input and output constraints. A multi-input multi-output (MIMO) CARIMA model identified through identification experiments is used to predict the future process outputs. Model parameters are estimated on-line to handle a time-varying nature of the process.

An analytical solution of a constrained MIMO GPC subject to input and output constraints is derived by solving a quadratic programming (QP) problem. The computation required by the analytical solution is substantially lower than that required by an algorithmic solution. For general cases of constrained MIMO GPC, an optimal solution is derived by solving a mixed-weights least-squares (MWLS) problem. In the use of MWLS, a control performance index can be easily augmented and the weights can be modified in a manner that encompasses both the requirements for the future control movements to lie inside the feasible region and to minimize the control performance index. If the constrained optimization problem is unfeasible, the MWLS will converge to the point that minimizes the maximum constraint violation. The proposed control schemes were
tested on the simulation model developed using the mechanistic and empirical methods to describe the behavior of a real process. Simulations demonstrated the proposed control schemes’ efficiency and capability of handling problems in one algorithm.

A linear model-based control strategy may lead to system oscillation or even instability if a refiner in the process is operated near maximum load point because at this point the nonlinearity between refiner motor load and plate gap becomes severe. To overcome the problem, a nonlinear Laguerre model - a type of orthonormal functions - is used to represent the nonlinear relationship. A MIMO Laguerre model-based predictive controller is then derived as an alternative for the control of a wood chip refiner. The Laguerre model can be arranged in linear form in model parameters so that the standard recursive least squares (RLS) can be used for on-line parameter estimation according to which controller parameters are adjusted. Simulations demonstrated that in the use of the Laguerre model-based control scheme, the nonlinearity of the process can be represented appropriately and the plate clashes resulting from the severe nonlinearity can be avoided. In addition, in the use of the Laguerre function representation the dynamics of an actual process can be described appropriately without the need for assumptions about the plant order and the time delay, i.e., accurate assumptions about their true values are not necessary.
# Table of Contents

Abstract ii  

List of Tables ix  

List of Figures x  

List of Symbols xiv  

Acknowledgment xvi  

1 Introduction 1  

1.1 General Introduction 1  

1.2 Background and Literature Review 2  

1.2.1 Mechanical Pulping Processes 2  

1.2.2 Refining Theory 6  

1.2.3 Control of A Refiner Mechanical Pulping Process 8  

1.2.4 Control Problems 14  

1.3 Objective of this Study 16  

1.4 Contributions of this Thesis 17  

1.5 Outline of the Thesis 20  

2 Process Variables, Modelling and Simulations 22  

2.1 Introduction 22  

2.2 Process Variables 23
3.5.3 Result Discussions ........................................... 76
3.6 Summary and Conclusions .................................... 77

4 Constrained Multivariable Model-Based Predictive Control 78
  4.1 Introduction .................................................. 78
  4.2 Unconstrained Generalized Predictive Control ............. 81
    4.2.1 Basic SISO Generalized Predictive Control ............ 81
    4.2.2 Basic MIMO Generalized Predictive Control .......... 87
  4.3 Recursive Parameter Estimation in Adaptive Control ....... 90
  4.4 Control under Constraints ................................ 91
  4.5 Constrained Control via Quadratic Programming .......... 95
    4.5.1 Quadratic Programming via Kuhn Tucker Multipliers Computation 96
    4.5.2 Analytical Solution for Constrained SISO GPC .......... 98
    4.5.3 Analytical Solution for Constrained MIMO GPC ....... 99
  4.6 Constrained Control via Mixed-Weights Least-Squares Algorithm ..... 102
    4.6.1 Preliminary ........................................... 104
    4.6.2 Constrained Control via Mixed-Weights Least-Squares .... 107
    4.6.3 Control Law ........................................... 110
  4.7 Simulations .................................................. 112
  4.8 Summary and Conclusions .................................. 122

5 Laguerre Function-Based Adaptive-Predictive Control 123
  5.1 Introduction .................................................. 123
  5.2 The Laguerre Functions .................................... 126
    5.2.1 The Linear Laguerre Functions ....................... 126
    5.2.2 The Nonlinear Laguerre Functions .................... 129
  5.3 SISO Laguerre Function-Based Predictive Control ........ 131
7.2 Future Work ......................................................... 192

Bibliography ......................................................... 194

Appendices .......................................................... 204

A Time Series Analysis Tools ..................................... 204
  A.1 Auto-Correlation Functions .................................. 204
  A.2 Partial Auto-Correlation Functions ......................... 205
  A.3 Cross-Correlation Functions .................................. 206

B Optimal Solution via Analytical Quadratic Programming ... 208
  B.1 Constraints ...................................................... 208
  B.2 Optimal Solution via Analytical QP .......................... 210

C Constraint Mapping .............................................. 218
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Primary Refining Conditions</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>First Stage Refining at $F_d = 4.7\text{kg/sec}$</td>
<td>43</td>
</tr>
<tr>
<td>2.3</td>
<td>First Stage Refining at $E = 3.4\text{GJ/t}$</td>
<td>45</td>
</tr>
<tr>
<td>2.4</td>
<td>Effects of First Stage Refining on Pulp Properties, $E=3.4\text{GJ/t}$</td>
<td>48</td>
</tr>
<tr>
<td>2.5</td>
<td>Secondary Refining Conditions</td>
<td>49</td>
</tr>
<tr>
<td>2.6</td>
<td>Effects of Secondary Stage Refining at $E=2.5\text{GJ/t}$</td>
<td>53</td>
</tr>
<tr>
<td>2.7</td>
<td>Effects of Secondary Stage Refining on Pulp Properties, $E=2.5\text{GJ/t}$</td>
<td>53</td>
</tr>
<tr>
<td>3.8</td>
<td>ML/CP Parameter Estimates, Estimated Standard Deviations and FPE</td>
<td>64</td>
</tr>
<tr>
<td>3.9</td>
<td>ML/TS Parameter Estimates, Estimated Standard Deviations and FPE</td>
<td>68</td>
</tr>
<tr>
<td>3.10</td>
<td>ML/DW Parameter Estimates, Estimated Standard Deviations and FPE</td>
<td>72</td>
</tr>
<tr>
<td>3.11</td>
<td>Identification Results</td>
<td>76</td>
</tr>
<tr>
<td>4.12</td>
<td>Kuhn-Tucker Multipliers for Different Cases</td>
<td>98</td>
</tr>
</tbody>
</table>
List of Figures

1.1 A simplified version of two-stage wood chip refining process 5
1.2 A double-disc refiner 6
1.3 Plate surface of a typical primary refiner 7

2.4 Interactions between manipulated variables, operating variables, pulp quality variables and disturbances for a two-stage refining 29
2.5 Mass and energy flows in the 1st-stage refiner 32
2.6 Model structure of the 1st-stage wood-chip refining 41
2.7 The 1st stage refining at $F_d=4.7$ kg/sec: effects of closing pressure $P_c$ and chip screw speed $S_c$ (--- $S_c=25$ rpm ($F_c=271$ t/d), --- $S_c=27.5$ rpm ($F_c=298$ t/d) and -- $S_c=30$ ($F_c=325$ t/d)) 44
2.8 The 1st stage refining at $E=3.4$ GJ/t: effects of dilution flow rate $F_d$ and chip screw speed $S_c$ (--- $S_c=25$ rpm ($F_c=271$ t/d), --- $S_c=27.5$ rpm ($F_c=298$ t/d) and -- $S_c=30$ ($F_c=325$ t/d)) 45
2.9 The 1st stage refining at $E=3.4$ GJ/t: effects of rotational speed $w$ and inlet consistency $c_i$ (--- $c_i=20\%$, --- $c_i=23.5\%$ and -- $c_i=27\%$) 47
2.10 The 2nd stage refining at $E_2=2.5$ GJ/t: effects of the rotational speed $w$ and the inlet consistency $c_i$ (--- $c_i=26\%$, --- $c_i=30.5\%$ and -- $c_i=35\%$) 51
2.11 Two-stage refining at $E_2=2.5$ GJ/t: effects of the 2nd-stage rotational speed $w$ and the inlet consistency $c_i$ (--- $c_i=26\%$, --- $c_i=30.5\%$, and -- $c_i=35\%$) 52
3.12 Input to and output from a dynamic process 56
3.13 Motor Load (ML) response to Closing Pressure (CP) ........................................ 65
3.14 ML/CP impulse response estimate ................................................................. 66
3.15 ML/CP residual correlations (d=0) ............................................................... 66
3.16 ML/CP residual correlations (d=1) ............................................................... 67
3.17 CP/ML input, output and model prediction ..................................................... 67
3.18 Motor Load (ML) response to Transfer Screw Speed (TS) ................................. 69
3.19 ML/TS impulse response estimate ................................................................. 70
3.20 ML/TS residual correlations .......................................................................... 70
3.21 ML/TS residual correlations (d=1) ............................................................... 71
3.22 ML/TS input, output and model prediction ..................................................... 71
3.23 Changes in Valve Positions (VP), Dilution Water (DW) and Motor Load (ML) ................................................................. 73
3.24 Motor Load (ML) response to total Dilution Water (DW) ................................. 74
3.25 ML/DW impulse response estimates ............................................................... 74
3.26 ML/DW residual auto- and cross-correlations ................................................ 75
3.27 ML/DW residual auto- and cross-correlations (d=1) ........................................ 75
3.28 Dynamic and static relationships between inputs and outputs ......................... 77

4.29 Basic structure of constrained model-based predictive control ....................... 95
4.30 Constrained GPC via analytical QP (t ≥ 300 sampling intervals: -1.5 ≤ u_t ≤ 1.5,
                          -2.0 ≤ Δu_t ≤ 2.0) ................................................................. 113
4.31 Constrained GPC via MWLS (t ≥ 300 sampling intervals: -1.5 ≤ u_t ≤ 1.5,
                          -2.0 ≤ Δu_t ≤ 2.0) ................................................................. 114
4.32 Criterion for stopping the iteration .......................................................... 116
4.33 Constrained GPC via analytical QP (t ≥ 300 sampling intervals : -4.0 ≤ u_1 ≤ 4.0,
                          -4.0 ≤ u_2 ≤ 4.0, |Δ u_1| ≤ 4.0, |Δ u_2| ≤ 4.0) ....................... 118
4.34 Constrained GPC via MWLS (t ≥ 300 sampling intervals: -4.0 ≤ u₁ ≤ 4.0, -4.0 ≤ u₂ ≤ 4.0, |Δ u₁| ≤ 4.0, |Δ u₂| ≤ 4.0) ........................................... 119
4.35 Output constrained GPC via analytical QP (t ≥ 300 sampling intervals: -1.0 ≤ y₁ ≤ 1.0) ................................................................. 120
4.36 Output constrained GPC via MWLS (t ≥ 300 sampling intervals: -1.0 ≤ y₁ ≤ 1.0, -1.0 ≤ y₂ ≤ 1.0) .................................................. 121

5.37 Laguerre ladder networks ................................................................. 127
5.38 The discrete-time nonlinear Laguerre network ........................................ 131
5.39 Linear model-based control will lead to oscillation when the nonlinearity
is represented by the dashed curve [1] ................................................ 135

6.40 Specific energy is the area under curve of N plotted against I ............... 145
6.41 Block diagram for the control of a two-stage TMP plant .......................... 149
6.42 Constrained multivariable adaptive-predictive control of the primary refiner 158
6.43 Input-output dynamic and static relationship for the primary refiner ......... 166
6.44 Constrained control via MWLS, t ≥ 200 sec: |Δ u₁| ≤ 50, |Δ u₂| ≤ 0.2,
|Δ u₃| ≤ 0.5 (q₁ = q₂ = 1, ρ₂ = 1.0, ρ₁ = ρ₃ = 0.0001) ........................................ 168
6.45 Constrained control via analytical QP, t ≥ 200 sec: 820 ≤ u₁ ≤ 1150,
3.4 ≤ u₃ ≤ 5.2 (q₁ = q₂ = 1, ρ₂ = 1.0, ρ₁ = ρ₃ = 0.001) ................................ 169
6.46 Unconstrained controller (q₁ = q₂ = 1, ρ₂ = 1.0, ρ₁ = ρ₃ = 0.0001) ........... 171
6.47 Unconstrained controller (q₁ = q₂ = 1, ρ₂ = 10, ρ₁ = ρ₃ = 0.0001) .......... 172
6.48 Unconstrained controller (ρ₂ = 1, ρ₁ = ρ₃ = 0.0001) ............................ 175
6.49 Unconstrained controller (ρ₂ = 1, ρ₁ = ρ₃ = 0.0001) ............................ 176
6.50 The gain between motor load and plate gap is subject to slow drift and a
sudden change in the sign from negative to positive .................................. 177
6.51 Structure of a 2×2 Laguerre model for a TMP refiner ............................ 180
6.52 Motor load setpoint unreachable ........................................ 185
6.53 A plate clash was avoided by the proposed control scheme ........ 186
6.54 Response to the changes in consistency setpoint ....................... 187
List of Symbols

The variables appearing in Section 2.3 are given in the follows:

\( Q \) : the volumetric flow rate of inlet wood chips (dry wood + moisture) \((m^3/s)\)

\( v \) : the chip velocity in chip transfer screw direction \((m/s)\)

\( S_c \) : the chip screw rotational speed (revolutions/min or rpm)

\( \rho_c \) : the chip bulk density of incoming wood chips \((kg/m^3)\)

\( f_c \) : the wood chip flow rate at inlet of a refiner \((kg/s)\)

\( s \) : the solid content of incoming chips \((\%)\)

\( e_c \) : the energy in inlet chips \((kJ/s)\)

\( e_d \) : the energy in dilution water \((kJ/s)\)

\( e_{sw} \) : the energy in the seal water \((kJ/s)\)

\( P \) : the net power drawn by a motor \((kJ/s)\)

\( \eta \) : the power efficiency coefficient \((\%)\)

\( e_{s1} \) : the energy in back flow steam \((kJ/s)\)

\( e_{s2} \) : the energy in forward steam \((kJ/s)\)

\( e_p \) : the energy in outflow pulp \((kJ/s)\)

\( F_c \) : the dry chip flow rate to a refiner \((kg/s)\)

\( f_f \) : the dry fiber flow rate in outlet pulp \((kg/s)\)

\( f_{w1} \) : the water flow rate in inlet chips \((kg/s)\)

\( f_{w2} \) : the water flow rate in outlet pulp \((kg/s)\)

\( F_d \) : the dilution water flow rate \((kg/s)\)

\( F_{sw} \) : the seal water flow rate \((kg/s)\)
$T_c$ : the inlet chip temperature(°C)
$T_p$ : the outlet pulp temperature(°C)
$T_d$ : the dilution water temperature(°C)
$T_{sw}$ : the seal water temperature(°C)
$c_c$ : the specific heat of dry chips($kJ/kg.°C$)
$c_f$ : the specific heat of dry fiber($kJ/kg.°C$)
$c_w$ : the average specific heat of water($20° - 150°C$)($kJ/kg.°C$)
$L$ : the enthalpy of steam($kJ/kg$)
$E$ : the specific energy($kJ/kg$)
$\tau$ : the residence time(s)
$w$ : the refiner disc rotational speed(radian/s)
$c_i$ : the inlet consistency(%)n
$r_1$ : the inlet radii of refining zone(m)
$r_2$ : the outlet radii of refining zone(m)
$L_s$ : the latent heat of steam($kJ/kg$)
$N$ : the number of refiner bar impacts received by unit pulp
$I$ : the average specific energy per refiner bar impact or refining intensity($kJ/kg$)
$\dot{e}$ : the specific refining power($kJ/kg/s$)
Acknowledgment

My greatest thanks go to my senior supervisor, Prof. Guy A. Dumont for his invaluable guidance and advice in my progress throughout the research, for his steady assistance, supervision and support in carrying out this study, for his continuous encouragement and help in my Ph.D. study, and for his valuable time and attention he devoted to me.

I would also like to thank Dr. Patrick Tessier, Dr. Bruce Allison and Dr. Ye Fu for their stimulating discussion and helpful suggestions in my thesis work.

I wish to express my gratitude to the members of my supervisory committee, Prof. Guy A. Dumont, Prof. Michael Davies and Prof. Ezra Kwok for their valuable discussions and assistance which proved to be invaluable for me to finish my thesis.

I thank all the members of Control Group from the UBC Pulp and Paper Centre for their valuable inputs in my thesis work during my stay at UBC.

Finally the financial support from Dr. Guy A. Dumont and NSERC Wood-Pulps Network Centres of Excellence is gratefully acknowledged.
To my husband, parents and brothers

Huaijing, with love
Chapter 1

Introduction

1.1 General Introduction

In the pulp and paper industry, wood pulping refers to the process by which wood chips are reduced to individual fibres. This task can be accomplished chemically (chemical pulping) or mechanically (mechanical pulping) or by combining these two methods. Each of them produces substantially different fibre characteristics.

In chemical pulping, lignin which binds the wood fibres together is dissolved chemically. The resultant fibres are longer, more flexible and considerably stronger than the mechanical fibres. During the manufacturing of chemical pulp, most of the lignin is successfully removed. Components other than lignins, such as hemicellulose and cellulose, are removed simultaneously, hence the yield of chemical pulping is lower (40-45 %) than the yield of mechanical pulping (up to 95 %) [2].

In mechanical pulping, the lignin is fractured out and not removed. Therefore, more wood constituents, such as fibre bundles, damaged fibres and fibre fragments with some whole fibres, are retained. This results in a higher yield (up to 95 %) in converting the wood into pulp [2]. Because of the mixed nature of the particles in the pulp, mechanical pulp has high light scattering coefficients as well as high bulk and relative stiffness [3]. However, mechanical pulp has small average particle length. The resulting paper consequently has low strength and high bulk, and discolors easily. In order to produce pulp
with the physical properties desired for higher quality printing, mechanical pulping combined with some chemical treatment can provide an alternative to mechanical pulping. The major drawback of mechanical pulping is that a greater amount of power is required than in chemical pulping. Hence equipment utilization, pulp quality level and product uniformity must be increased. This can only be accomplished by modifying the process, and by improved equipment and process control [3]. This thesis partially addresses the problems. In particular, the thesis attempts to develop a suitable control strategy for a two-stage thermo-mechanical pulping (TMP) plant.

The purpose of this chapter is to introduce the problems and state the thesis objectives. Section 1.2 includes the background material relevant to mechanical pulping processes and their control. Section 1.2.1 gives a brief introduction to mechanical pulping processes including unit operations; Section 1.2.2 describes refining theories; Section 1.2.3 briefly reviews the past and existing control schemes; and Section 1.2.4 states some existing control problems. Thesis objectives, contributions and outline are given in Sections 1.3, 1.4 and 1.5.

1.2 Background and Literature Review

1.2.1 Mechanical Pulping Processes

Depending on the machinery and action utilized, mechanical pulps can be produced by two different processes: grinding and refining. Grinding or stone groundwood is the oldest method, in which wood logs are forced against a rapidly revolving roughened grindstone and converted to individual fibres. In refining, wood chips are fed between two metal discs (at least one of which rotates) of a refiner and converted to individual fibres. These two processes result in significantly different pulp characteristics. Groundwood pulp has a higher content of fine material due to the abrasive action, whereas refiner
Chapter 1. Introduction

pulp has a smaller content of fine material [4] but a higher content of long fibres. As a result, refiner pulping produces much stronger fibres than stone groundwood. Various types of refiner mechanical pulp can be obtained by modifying a refiner pulping process. Thermo-mechanical pulping (TMP) is a modification of a refiner mechanical pulping (RMP) process. It involves steaming the raw material for a short period of time prior to refining to soften the chips. The resulting pulp has a greater percentage of long fibres and less shives than those of an RMP. Chemi-thermo-mechanical pulping (CTMP) is a modification of a TMP with some chemical treatment in order to obtain pulp with desired physical and chemical properties for higher grade applications. Due to the thermal pretreatment to the chips and chemical action in separation of fibres, CTMP produces pulp with a higher content of longer fibres and a higher brightness.

Unit Operation

A thermo-mechanical pulping process generally involves three main operation areas:

1. wood chip pretreatment

2. wood chip refining

3. pulp processing

*Wood chip pretreatment* consists of chip screening to remove under or oversize material; chip washing to remove rocks, metal and sand; chip steaming to soften lignins binding the fibres so that produced pulps have a greater percentage of long fibres and less shives. In chemi-thermo-mechanical pulping, the wood chips are impregnated with chemicals to improve pulp brightness and strength.

*Wood chip refining* aims at breaking chips into individual fibres (see following description).
Chapter 1. Introduction

Pulp processing is aimed at enhancing and controlling pulp quality. It consists of: latency removal to straighten the fibres; pulp screening and reject refining to remove unrefined fibre bundles; pulp cleaning to remove heavy contaminants; pulp washing to remove wood resins and metallic ions; and pulp bleaching to increase brightness.

Wood Chip Refining

Generally, wood chip refining is carried out in two stages as shown in Figure 1.1. Wood chips (dry wood and water moisture) are introduced into the open eye of the primary refiner by the chip transfer screw. As the wood chips pass through the refining zone between the rotating plates (or discs), they are progressively broken down into smaller particles and finally into individual fibres. In the secondary refiner, the fibres are further refined according to the pulp quality requirements. Simultaneously, dilution water is added to both primary and secondary refiners to control the consistencies in the refining zones. Plate gap for each refiner is controlled by either an electro-mechanical or hydraulic loading system mounted in the refiner. Following each of the refiners, cyclones are installed to separate steam from the produced pulp and to cool the hot pulp to prevent brightness loss.

A wood chip refiner, a key element in the pulping process, is either a single-disc (one rotating disc and one stationary disc), a double-disc (two discs rotating in opposite directions, Figure 1.2) or a twin (a rotating double-sided disc between two stationary discs) configuration. A refiner disc, driven by a large motor, usually rotates at 1200 – 1800 r/min. Refiners are currently available with up to 70 inch (1.8m) diameter and 18,000 HP (13,400 kW) of power supplied to each plate [2]. A large-diameter high-powered refiner is generally developed for large-capacity refining. Plate gap is accurately controlled to about 0.1 mm and is always in the range of 0.05 to 0.2 mm [5]. Refiner plates are generally divided into three sections: the breaker bar section, the intermediate
Chapter 1. Introduction

Figure 1.1: A simplified version of two-stage wood chip refining process

bar section and the fine bar section (Figure 1.3). The breaker bar section (closest to
the eye of the refiner) has wide bars with deep grooves, serving to break up the chips
as they enter the refiner. The intermediate and fine bar sections consist of progressively
narrower bars and shallower grooves, which serve to refine continuously the partly refined
pulp. The three sections may vary in size depending on the refining stage. Plate life is
generally in the range 500 - 1000 hours and varies with different material and refining
stage. The plate life of a primary refiner is longer than that of a secondary refiner because
refining action in the secondary refiner is greater than that in the primary refiner. In the
refining zone (or the fine bar section), most refining is done at consistencies in the range
of 16% to 50% [2]. Most of the energy introduced by a refiner is converted into the heat
which produces hot water and steam. The steam produced in the refining zone develops
pressure substantially above those at either the inlet or the discharge. This amount of
steam can be indicated by the temperature measurement in the refining zone. The peak
temperature may range from 115 °C to 145 °C, equivalent to a peak pressure in the order
of 170 kPa to 446 kPa [6]. Additional information on refiner mechanical pulping can be
found in [2, 4, 7, 8].
1.2.2 Refining Theory

Intensive theoretical studies have been made by many researchers towards developing a unified theory of refining. Much work done in the past was based on the product quality-energy consumption relationship, i.e., pulp quality such as freeness was determined by the energy applied. However, recent work showed that specific energy alone does not fully determine pulp properties and refining action [9, 10, 11, 12, 13]. Many researchers agree that pulp quality is determined by two basic variables: the energy per unit mass of pulp or specific energy \((E)\) and the specific energy per bar impact or refining intensity \((J)\), rather than by the specific energy alone [5, 14, 3, 15, 9, 10, 11, 12, 13]. For a given refiner design and a constant disc rotational speed, the refining intensity is largely dependent on refining consistency while it is insensitive to the specific energy [16].

The influence of the refining intensity on pulp quality in low consistency refining has been known for a long time [10], but it has only recently been pointed out by Lunan et
that it may also be important in determining the pulp quality in high consistency refining. Lunan’s idea has been further investigated by Miles [10], which showed that for a given specific energy under a wide variety of operating conditions, the refining intensity was strongly correlated to the pulp quality. The work by Miles et al. [10, 11] and Rodarmel [18] also showed that for a given specific energy, low intensity refining produces pulps of higher strength but lower opacity, while high intensity refining results in pulps of lower strength but better printing properties. It was suggested by Miles et al. recently that at a given total specific energy, the combination of high intensity-low energy refining in the first stage followed by low intensity-high energy refining in the second stage could produce pulp with improved printing properties with no loss in strength [11].

Despite significant progress by many researchers last few decades, the refining action in the refining zone is still not well understood. To optimize the refining processes, more theoretical and experimental work on the mechanics and dynamics of the fibres between two plates of a refiner are needed.
1.2.3 Control of A Refiner Mechanical Pulping Process

The control of mechanical pulping processes has been in a state of ongoing development since the mid-1970s [19]. Significant progress has been made towards refining optimization, cost reduction and product quality improvement. The following briefly reviews some control strategies and applications available in the literature. Additional information on the control of refiner mechanical pulping processes can be found in [20, 21, 22, 23, 24, 25, 26, 27, 28, 29].

Past Control Strategies

Specific Energy Control

Specific energy is a main variable affecting pulp properties. A traditional pulp quality control strategy is by manipulating specific energy based on the quality-energy relationship. Specific energy is defined as the ratio of net power or motor load applied to dry fibre mass throughput or production rate. Therefore, the target of specific energy is often achieved through the control of production rate or motor load. Many mills use chip screw speed to set the desired production rate. However, the production rate may still vary due to changes in the quality of raw material such as chip bulk density and moisture content.

Motor Load Control

The motor load is most sensitive to plate gap, and therefore one way of setting and maintaining the motor load is to adjust plate gap. Various control approaches [27, 30, 31, 32, 33] have been proposed to solve motor load control problems. With these control approaches, the motor load was controlled through manipulating the plate gap while the throughput was kept constant and the refining consistency was not under control. Although load controls via gap adjustment will reduce load variations, the motor load can still vary due to the variations in raw material quality such as chip size,
Chapter 1. Introduction

moisture and density. To deal with the fluctuations in raw material quality, an alternative has been proposed [34, 35, 36, 37], where the motor load was controlled by automatically adjusting chip feed indicated by the chip screw speed. Using this control approach, the lower frequency load fluctuations induced by wood quality variations can be reduced, but not the higher frequency fluctuations in the motor load [34]. Therefore, further work needs to be done to reduce the higher frequency fluctuations in the motor load.

Consistency Control

Refining zone consistency is known to have a major effect on pulp properties at given specific energy. An improved consistency control will significantly enhance the refiner performance by eliminating the unavoidable feeding disturbances [38], while a poor consistency control can lead to unstable refiner operation [28]. Various control strategies have been proposed with the help of on-line consistency sensors [39, 38, 40, 41, 35, 36], where the consistency at the blow-line was controlled by manipulating the dilution flow at the inlet. Improved consistency control can provide a significant reduction in the fluctuation of the motor load and obtain large improvements in stabilizing TMP/CTMP plants, thus resulting in energy savings, control efficiency and improved pulp quality [41, 40].

Pulp Quality Control

The pulp quality can be defined by many property variables, but in many cases freeness is still seen as an important variable to characterize the strength of mechanical pulp [42]. Various control strategies have been proposed [26, 27, 43, 44, 42, 45] in the area of freeness control. The objective of these works was to control the freeness to the quality target by adjusting the specific energy while minimizing freeness deviation from the target as well. Freeness was controlled by adjusting the specific energy setpoint through manipulating the motor load or the production rate. Simulation studies and industrial applications showed that the proposed controls brought about freeness variation reduction
and some improvement in pulp quality. However, measuring and controlling pulp quality by freeness alone is not satisfactory because pulp quality is determined by at least two property variables, rather than by freeness alone. An alternative to the single-variable control has been proposed [19], where both freeness and long fibre content were controlled by selecting the optimal setpoints of a TMP operation through a model-based analysis system. It showed the potential for multivariable quality control of the refining operation, but one of the weaknesses of this approach was maintaining accurate information flow to the model [19].

**Advanced Control Applications**

Proportional-Integral-Derivative (PID) control is the most prevalent form of industrial control. However, standard PID techniques are not suitable for the control of the processes with time delay, nonlinearity, multivariable and time-varying gain/dynamics. Alternatively, advanced control methods will be able to handle the problems more efficiently and help to ensure optimum control performance. Following is a brief review of some advanced controls available in the literature. Additional information in advanced control applications can be found in [21, 46, 23, 47].

*Adaptive Control*

Motor load is most sensitive to plate gap adjustments. However, control of the motor load by adjusting the plate gap is difficult due to the fact that the gain between the load and the gap is subject to a slow drift as well as a sudden change in sign. To deal with the problem, various control schemes have been proposed [24, 27, 30, 48]. Dumont first applied an adaptive control scheme to the problem [30], where a Dahlin regulator was tuned using the gain estimate through a recursive least-squares estimator with a variable forgetting factor. Industrial trials showed that the proposed regulator was able to track the slow gain drift and a sudden change in its sign [30, 43]. However, one problem with
this technique is the tedious tuning of the forgetting factor algorithm parameters, so it may be unreliable in a continuous usage environment [32].

**Dual Adaptive Control**

To improve the reliability of the adaptive controller proposed in [30], Dumont and Åström [32] proposed a dual control strategy, where additional terms reflecting the non-linear nature of the process were introduced into a performance index, rather than using an index function reflecting the output error only. In the case of an unreachable motor load setpoint, the controller will tend to keep the gain negative instead of eliminating the output error. The proposed dual control strategy was further modified with an improved criterion by Allison *et al.*[48]. Preliminary results of an industrial refiner control implementation showed the general success. However, one problem with this technique was the computational difficulty attached to solving the dual control problem on-line [31].

**Adaptive Inferential Control**

An adaptive inferential control strategy was proposed by Kooi [49] for the closed loop control of freeness through manipulating the gap. The control strategy was for the cases when the measurement of freeness is not readily available either due to lack of a reliable on-line sensor or the long sampling time to obtain the test result. With this strategy, rapidly measured specific energy was used to infer the freeness based on a dynamic-noise model which linked the plate gap and the specific energy to the freeness. Due to the time-varying and stochastic natures of the process, the controller parameters were tuned based on on-line estimates of the model parameters. However, this single-variable control approach was limited by its definition of pulp quality.

**Minimax Robust LQ Control**

A linear quadratic (LQ) control based on multimodelling techniques was proposed by Toivonen and Tamminen [45] for the control of freeness in a TMP plant. The process was assumed to be described by a set of linear models presenting different operating
Chapter 1. Introduction

situations. The control problem was then reduced to a single constant control law so that satisfaction of control performance in terms of the output error was obtained for all set of the models. The idea was achieved by formulating a minimax control problem, in which the maximum output error of a given sets of models was minimized subject to a constraint on the maximum input variance. The proposed control strategy was tested in a TMP plant for the control of freeness through manipulating the chip feed. The results showed that the performance of the proposed controller was better than that of a manual control or the control by a PI-regulator [45]. However, this approach is only suitable when the range of the parameter changes can be estimated and when the changes are not too large.

Neural Networks

In general, the objective of a neural network based control approach is to develop an algorithm to adjust the parameters of the network based on a given set of inputs and outputs. The error between the desired output and the computed output at the output layer of the network is back propagated to the hidden and input layer for weight updating. The use of neural networks with back-propagation learning was proposed by Kooi and Khorasani [50] for the control of the specific energy in a TMP plant. The proposed controller consisted of one processing element in each input and output layer, and two hidden layers with 10 processing elements in each hidden layer. It was reported in [50] that this neural network-based controller was robust in the sense that it needed no prior knowledge of process order, time delay, dynamics and noise. However, the use of a neural network model usually requires a large number of weights and the weights are not in a linearized form, thus the estimation converges slowly [51]. The performance index via model parameters also becomes complicated. These characteristics are the main difficulties when applying a neural network model in adaptive control.

Laguerre Function-Based Control
A predictive control strategy based on a nonlinear Laguerre model was introduced by Fu and Dumont [51] to handle the nonlinear relationship between refiner motor load and plate gap. The controller parameters were tuned based on the estimates of the Laguerre model parameters. Simulation results showed that in the case of an unreachable motor load setpoint, the proposed controller was always trying to keep the gain negative instead of eliminating the output error [51]. The above single-variable control idea was later extended into a $2 \times 2$ system by Du et al.[52] to overcome the limitation with a single-variable control philosophy. With this control approach, the motor load and the outlet consistency were controlled by adjusting the gap through manipulating the closing pressure, and the dilution flow rate. The process model was based on a mixed linear-nonlinear Laguerre function representation, where the nonlinear Laguerre functions were used to model the gap-load nonlinear relationship whereas the linear Laguerre functions were used to model the dilution-consistency relationship. The predictive controller parameters were tuned based on on-line estimates of the model parameters. Simulation results showed that the use of this approach was not only able to handle the nonlinearity between the load and the gap, but also able to control the refining consistency to its setpoint according to the requirement of pulp quality [52]. Also, the use of Laguerre function-based approaches can eliminate the requirement for the accurate knowledge of the model order, time delay and process dynamics.

*Constrained Predictive Control*

For the purpose of control applications, it is always necessary to take process constraints into account in a controller design, due to safety, equipment physical limitations and product quality requirement. The controller should not be allowed to drive input-output variables outside their specified bounds. Du and Dumont [53] first proposed a constrained control approach for the control of a wood chip refiner. In the use of this approach, a wood chip refiner was modelled as a $3 \times 2$ system. A constrained predictive
control law based on GPC was derived by taking constraints into account. Optimum solution was calculated through quadratic programming. An adaptive control scheme was used to handle the time-varying and nonlinear natures of the process. The simulation results showed [53] that the proposed controller was able to maintain the outputs at their setpoints without violating the constraints. It was also shown that in the use of the predictive control strategy, input-output constraints can be handled explicitly and efficiently.

1.2.4 Control Problems

Extensive research effort towards better control of a TMP plant has been made over the last two decades. However, the control of today's refiner mechanical pulping plant is still relatively primitive [33]. This may be due to three major problems: (i) a lack of on-line sensors for measuring key process variables, (ii) unknown process mechanism and (iii) limitations of the past control methods.

A lack of sensors

On-line sensors are the key to successful control of industrial processes, but they are not available sometimes. Sensors required for the control of refiner mechanical pulping can generally be divided into three groups: sensors for raw material, sensors for refining conditions and sensors for pulp qualities.

The main difficulty in the control of specific energy is our inability to measure fibre feed rate or throughput accurately. For a given chip transfer screw speed, throughput variations are mainly resulted from changes in inlet chip bulk density and moisture content. Even when throughput is constant, pulp quality may still vary because of changes in wood species. Different wood species will require different energy input due to differences in fibre length, coarseness, etc. The variations in raw material should be compensated in order to produce pulp with the desired quality. However, due to a lack of on-line sensors
neither these variations nor their effects on the pulp quality can be measured accurate and fast enough at present [39, 54].

Sensors for refining conditions are also required to help stabilize refiner operation and compensate for immeasurable changes. Consistency in the refining zone is perhaps the most important variable in determining the quality of the pulp [55], but it is difficult to measure the consistency in high consistency refining. Presently, several blow-line consistency sensors are available [56, 41, 40] with varying degree of success in their applications. There has been no publication assessing the performance of these sensors. This may be partially due to the fact that these sensors are all relatively new and their reliability needs to be investigated. Refining intensity is also an important variable in determining the quality of the pulp for a given specific energy. If there was a technique available for on-line measuring the refining intensity, there would be one more degree of freedom in the control of refining and pulp quality.

On-line sensors for measuring pulp properties are absolutely necessary in manufacturing high quality pulp, but there is still a lack of fast, continuous, reliable on-line sensors for pulp quality [33]. A relatively sophisticated analyzer: the Pulp Quality Monitor or PQM [57, 25, 58] is currently available. It can be used for measuring pulp drainage, fibre length distribution and shive content. The PQM has the distinct advantage of being capable of performing several tests per hour in comparison to the relatively low frequency manual testing [59]. However, the PQM has not been widely accepted in North America because of its high cost and maintenance requirements [33]. Relatively simple and inexpensive quality analyzers are needed by many mills.

*Unknown Process Mechanism*

The main impediment to improving control of refiner mechanical pulping process is a lack of understanding of refining mechanisms. Over the last few decades, tremendous efforts have been made by many researchers and engineers towards a better understanding
Chapter 1. Introduction

of the wood chip refining process. Despite intensive activity in the field and improved knowledge of the refining process, there is still a lack of understanding of the refining mechanism in a disc refiner. More theoretical and experimental work are needed in the fields of the mechanism of energy transfer and mass flows in the refining zone during wood chip refining.

Limitations of Control Methods

A traditional method for the control of pulp quality is based on a quality-energy relationship. Associated controllers are single-input single-output linear model based with fixed controller parameters. In many cases, control actions are made and implemented on the assumption that the resulting control actions are feasible and can be implemented. However, an actual process is mainly a function of two variables: the input energy per unit of mass pulp or specific energy, and the specific energy per refiner bar impact or refining intensity. The process is characterized by nonlinearity, variable interactions, process constraints, time varying nature and external disturbances. Therefore, conventional control methods are not satisfactory and limited by their control performance.

1.3 Objective of this Study

In the past, various control strategies have been proposed for the control of wood-chip refining, but the associated control performances are limited for certain reasons. First, past control methods were more focused on controlling the process as a single-variable, time-invariant and linear process. They were not satisfactory because the real process is a stochastic one with multivariable, nonlinear and time-varying characteristics. Secondly, in many cases, past control approaches were developed based on a quality-energy relationship, i.e., pulp quality is determined by specific energy alone. However, pulp quality is a function of two variables: the specific energy and the rate of the transferred specific
energy or the refining intensity. For a given specific energy, pulp quality will vary as the refining intensity changes. Thirdly, the control calculations in past control applications were made and implemented with the assumption that the resulted control actions were feasible and could be implemented. This would limit the control performance because the inability to implement the control signal exists due to the reasons of safety, equipment physical limitations and product quality requirements. The limitations of the past control approaches motivated the current study.

The objective of this study was to develop a control strategy suitable for the control of an industrial TMP plant. To achieve this goal, three tasks were completed. The first of these was to model the process using mechanistic and empirical methods. The second task was to develop a constrained multivariable control strategy for the control of the TMP plant. The third was to investigate the proposed control strategy on a simulation model developed using the mechanistic and empirical models to approximate the characteristics of an industrial situation.

1.4 Contributions of this Thesis

This thesis describes the development of a novel control strategy for the control of a two-stage TMP plant. Desired pulp quality is maintained by selecting specific energy and refining intensity setpoints which are achieved through the control of production rate, motor load and refining consistency at each stage. The manipulating variables are hydraulic closing pressure on refiner plates and flowrates for both chip feed (at the 1st-stage) and dilution water. A constrained multivariable controller is developed based on the GPC algorithm because of its simplicity, flexibility and capability of handling problems in one algorithm. An optimal solution of the constrained MIMO GPC is obtained by solving a quadratic programming or mixed-weights least-squares problem. A MIMO
CARIMA model is used for the controller development. However, in the case of process severe nonlinearity, i.e., the incremental gain between refiner motor load and plate gap is subject to a sudden change in the sign, a mixed linear-nonlinear Laguerre model is used for the controller design.

The major contributions of this thesis are as follows:

- A steady state model is developed to link process input variables to output variables. The outlet consistency at each stage is modelled using mass and energy balances, showing that outlet consistency is a function of production rate, motor load and inlet consistency. If the measurements of refiner inlet and outlet consistency are available, the consistency model may be used to back calculate production rate.

- Process dynamics and disturbances are modelled through identification experiments performed on an industrial process. The process is identified as a multivariable dynamic-stochastic process with lower-order dynamics plus time delay. In the use of process identification (or the empirical method) rather than the mechanistic method, process dynamic modelling can be realized more easily without the need for the development of complex differential equations. The proposed control scheme will be based on the model structure obtained from process identification stage.

- A multivariable control strategy is developed for the control of the TMP plant. Desired pulp quality is achieved through selecting specific energy and refining intensity setpoints which are achieved through the control of production rate, motor load and outlet consistency at each stage. The manipulated variables are closing pressure on refiner plates and the flowrates for both chip feed and dilution water. The proposed control strategy will have more degrees of freedom in the control of the process, and the capabilities of optimizing process operation and improving product quality.
• A constrained multivariable controller for TMP is developed based on the GPC algorithm for the control of the process subject to input and/or output constraints. An optimal solution of the constrained controller is obtained through solving a quadratic programming (QP) problem. Simulations demonstrate that the control law given by input and output constraints is realizable. In the case of an unfeasible constrained control problem, an optimal solution is obtained by solving a mixed-weights least-squares (MWLS) problem and the MWLS will converge to a point that minimizes the maximum constraint violation. In addition, in the use of the MWLS the control performance index for any case of constrained control problems can be augmented easily and the weights can be modified in a manner that encompasses both the requirement for the future control movements to lie inside the feasible region and to minimize the control performance index. Computer simulations show that the performance of the constrained controller is better than the unconstrained one.

• For the case of the severe nonlinearity of the process, i.e., the incremental gain between refiner motor load and plate gap is subject to a sudden change in the sign, Laguerre functions – a type of orthonormal functions – are considered as an alternative to represent plant dynamics and nonlinearity. The multivariable predictive controller is then derived based on a mixed linear-nonlinear Laguerre model. The nonlinear Laguerre functions are used to represent the severe nonlinearity between refiner motor load and plate gap, while the linear Laguerre functions are used to approximate the relationships between other inputs and outputs. The major advantages of the Laguerre approximation are: (i) easy to model and reconstruct, (ii) flexible model structure and simple representation, (iii) understandable structure similar to transient signals, and (iv) the capability of representing a plant properly
without the need for assumptions about plant order and time delay. Computer simulations demonstrate the robustness and efficiency of the proposed control scheme, along with the ability to avoid plate clashes due to sudden changes in the sign of the incremental gain between refiner motor load and plate gap.

- The efficiency and applicability of the proposed control schemes are tested on a simulation model developed using a combination of the empirical and mechanistic methods to approximate an industrial situation. The simulation results demonstrate various advantages of the proposed control schemes with the potential to offer an automated control of mechanical pulp manufacturing.

1.5 Outline of the Thesis

After a brief review of the fundamentals related to refiner mechanical pulping processes, Chapter 2 describes static modelling of the process using the mechanistic method. A qualitative relationship between raw material quality, operating conditions and pulp property is presented. Simulations show the variable interactions and nonlinearity of the process. Chapter 3 presents dynamic modelling of the process using the empirical method. To avoid the complexity of deriving differential equations and the inability of modelling process disturbances through the mechanistic modelling, process identification is used to obtain the characteristics of process dynamics and disturbances. Chapter 4 describes the development of a constrained multivariable predictive controller for the control of a TMP plant. An optimal solution of the constrained MIMO controller for simple cases is derived by solving a quadratic programming problem, while for general cases an optimal solution is derived by solving a mixed-weights least-squares problem. Chapter 5 presents the development of MIMO Laguerre model-based controller to avoid system oscillation resulting from a linear-model-based controller in the case of severe nonlinearity of the
process. Using the Laguerre model-based control approach, the dynamics of the process can be represented appropriately without the need of assumption of the model order and time delay. The proposed control strategy for the control of the two-stage TMP plant is detailed in Chapter 6. Desired pulp quality is achieved through selecting specific energy and refining intensity setpoints which are obtained by changing production rate, motor load and outlet consistency. Simulation results demonstrate the efficiency and applicability of the proposed control schemes. The thesis conclusion and future work are given in Chapter 7.
Chapter 2

Process Variables, Modelling and Simulations

2.1 Introduction

A mathematical model, generally speaking, can be a completely mechanistic one, an empirical one or a combination of these two. A mechanistic model based on physical principles (or the first principles) is more desirable since it is generally applicable over a wide range of operating conditions and gives more insight into a process, but it may be time consuming to develop. By comparison, an empirical model derived by fitting a mathematical equation to a set of experimental data is easier to develop and possesses more simplicity, but because it approximates the process over a limited range of operation, it will present the real process with some limitation. Which approach is chosen for modelling depends on the purpose of the model. In the past, much work has been done on modelling refining process using different approaches. Most of the work focused on modelling the process in order to obtain more insight into the mechanism of the process, rather than for control purposes. The mathematical model developed in this study is for control purposes and will give a trade-off between the model accuracy and simplicity. The model is developed by using a combination of mechanistic and empirical methods. Such a mechanistic-empirical model not only gives some insight into the process mechanism, interactions, and nonlinearity, but also represents the characteristics of process dynamics and disturbances. Mechanistic modelling is given in the following, while empirical modelling will be presented in Chapter 3.
This chapter is organized as follows. Section 2.2 lists a number of main process variables which will be considered in process modelling, such as manipulated variables, operating variables, pulp quality variables and disturbance variables. Section 2.3 presents modelling of a two-stage refiner mechanical pulping process. The developed model will show process interactions and nonlinearity. Section 2.4 includes simulations of the model and result discussions. Summary and conclusions are given in Section 2.5.

2.2 Process Variables

Refining is a series of interrelated events designed to produce pulp. The interrelationships of these events can be expressed in terms of process variables. The variables, under this study, are divided into four groups: manipulated variables, operating variables, pulp quality variables and disturbance variables. Following are the key variables which have strong effects on wood chip refining. Other process variables such as steam flow, refiner plate pattern design, chip feed temperature, dilution water temperature, etc. can also affect chip refining, but they are not included either because the effects of the variables are small or because the variables are relatively constant during refining.

2.2.1 Key Manipulated Variables

Manipulated variables are the input variables which can be adjusted during the process operation to control the refining conditions. Changes in manipulated variables cause changes in the operating conditions and thereby changes in the end product quality.

*Chip Transfer Screw Speed*

The chip transfer screw, located prior to the primary refiner, is used to adjust the volumetric flow rate of wood chips to the process. Currently, many mills use transfer screw speed to set desired production rate or throughput. Any changes in the screw
speed will change dry fibre flow to the process. As a result, the energy input per unit dry fibre or specific energy will be varied.

*Hydraulic Closing Pressure*

The hydraulic closing pressure applied by a hydraulic loading system mounted in a refiner is adjusted to position the plate gap of the refiner. An increase (or decrease) in the hydraulic pressure results in a decrease (or increase) in the gap and hence an increase (or decrease) in motor load because motor load is more sensitive to the gap.

*Dilution Water Flow Rate*

The consistency in refiner refining zone is important to pulp quality. For a given specific energy, any change in the consistency may alter the rate of the specific energy input or refining intensity and as a result alter pulp quality. Also, excessive consistency variations can lead to unstable refiner operation (plate gap, power, etc.) [28]. Manipulating the dilution flow at the inlet of each refiner is the most convenient way of maintaining the consistency in the refining zone.

*Disc Rotating Speed*

Disc rotating speed is defined as the number of disc revolutions per minute. Any change in rotating speed will directly vary the centrifugal force on the pulp in the refining zone and thus change the residence time for a fibre to pass through the refining zone. Any change in residence time will alter the number of refiner bar impacts per unit pulp and for a given specific energy, this will in turn lead to changes in specific energy per bar impact or refining intensity. As a result, pulp quality will be changed. Generally, disc rotation speed is fixed once it is determined according to production requirement and refiner physical reliability. However, strong effect of disc rotating speed on the refining intensity suggests that manipulating disc rotating speed, at a given specific energy, will be a more sensitive way of adjusting pulp quality.
2.2.2 Key Operating Variables

The quality of the pulp produced depends very much on the operating conditions in the refining zone between plates. The operating conditions in the refining zone will change with changes in the manipulated variables.

*Motor Load*

A motor integrated with a refiner is a power actuator which delivers electrical energy to the refiner. Motor load or net power is defined as the electrical energy applied to the fibre flow in the refiner. At a given throughput, any change in motor load may vary pulp quality because motor load is proportional to specific energy. Motor load may vary with changes in throughput, plate gap, refining consistency, raw material quality, etc.

*Refining Consistency*

Refining consistency is important in a refiner mechanical pulping process, because it changes the energy-quality relationships [60]. At a given specific energy, refining at different consistencies results in the pulp with different quality [61]. Refining consistency may change with changes in throughput, dilution flow, motor load, etc. but the dilution flow is the most convenient way to be used to adjust the consistency.

*Plate Gap*

Plate gap is the separation of two discs of a refiner and is controlled by either a hydraulic or electro-mechanical loading system integrated into the refiner. The gap measurement can be obtained by a gap sensor in the refiner or indicated by the relative shaft position of the loading system [49]. In practice, the gap measurement is used for indicating the plate movement and/or interlocking purpose to prevent plate clashing, rather than for closed-loop refiner control. Any changes in plate gap will change the mechanical force exerted by the wood material in the refining zone on the plates, and as a result alert motor load and specific energy for a given throughput.
**Production Rate**

Production rate or throughput is defined as oven-dry \(^1\) wood flow fed to a refiner. Many mills use chip screw speed to set desired production rate. However, this is not accurate due to variations in chip density and moisture content. For a given motor load, any variations in throughput may result in changes in the energy input per unit fibre or specific energy and thus pulp quality.

**Specific Energy**

It is known that pulp quality is strongly affected by specific energy. Since specific energy is proportional to motor load and inversely proportional to throughput, the traditional method of adjusting pulp quality is through manipulating the motor load or the throughput.

**Refining Intensity**

At a given specific energy, refining intensity determines the rate of the energy inputs. For a given specific energy, different refining intensity will produce the pulp with quite different quality \([17, 62, 63, 10, 11]\). It is only in the last few years \([11]\) that the concept of refining intensity has been identified and its importance has been realized. It has been suggested \([64]\) that refining intensity must now be recognized as an important variable in all types of refining. Refining intensity is a function of specific energy, residence time, rotating speed and refiner design. However, for a given specific energy residence time is only available way to alert the refining intensity.

**Residence Time**

The residence time for a fibre to pass the refining zone is the most important parameter in the determination of the final pulp quality for a given specific energy \([10, 15]\). This is because changes in the residence time will bring about changes in the number of bar

---

\(^1\) oven-dry: condition of cellulosic material that has been dried to constant mass (or weight) at a temperature of about 105 °C.
impacts received by a fibre. As a result, average specific energy per impact or refining intensity will change if specific energy is constant. Residence time is a function of specific energy, inlet consistency, rotational speed and refiner design. For a given specific energy, manipulating inlet consistency is the most convenient way to alert the residence time.

2.2.3 Key Pulp Quality Variables

The pulp properties in terms of papermaking like freeness, fiber length, fiber specific surface, shive content, coarseness, flexibility, strength, etc. are important. The choice of different property variables depends on the different end users for the different pulp quality and also relies on the techniques available for obtaining the measurements of these variables. In the following, the Canadian Standard Freeness, long fibre content and shive content are considered as property variables. Such choice is not only because these variables can be used to predict handsheet strength and pulp drainage properties, but also because they can be measured using the techniques available on the market, such as Pulp Quality Monitor (PQM). PQM is an automated device for on-line measuring freeness, fibre length distribution and shive content. Pulp quality is dependent on the conditions of refining and can be consequently adjusted by changing the operating conditions through manipulating the manipulated variables.

**Canadian Standard Freeness**

The degree of refining done on pulp is measured by the drainage of pulp, called freeness. However, pulp drainability is usually obtained by means of the Canadian Standard Freeness (CSF) test, which is defined as the number of ml water collected from the side orifice of the standard tester when a dilute stock drains through a perforated plate under carefully controlled conditions [65]. Since freeness characterizes the strength of mechanical pulp, it is used to describe the strength of the produced pulp [42].

*Long Fibre Content*
Chapter 2. Process Variables, Modelling and Simulations

Long fibre content is another variable used to describe the strength of the produced pulp. The longer the individual fibre is, the stronger the pulp is. The long fiber content is defined as the percentage by weight of fibres retained on the 48-mesh screen of the Bauer-McNett fiber length classifier [66].

**Shive Content**

Shive content is defined as the percentage of oven-dry pulp retained on a standard slotted fractionating plate, usually with slit width of 0.15mm [65]. It is possible for pulps with identical long fibre content to have different strength and drainage properties due to the presence of the shives. So shive content is also used to indicate pulp strength and drainage properties.

2.2.4 Disturbance Variables

Disturbance variables are the variables which cannot normally be controlled and often cannot even be measured. The variations in wood quality and refiner plate conditions are the normal disturbances [67] influencing refining conditions and pulp quality.

**Raw material quality**

Very often, a major disturbance to a refiner mechanical pulping process is the variations in raw material quality, such as wood species, chip size, chip bulk density and chip moisture content. However, chip bulk density variation is a main cause of the disturbance [54]. Any change in the chip bulk density will change the flow rate of the dry wood mass or the throughput. For a given motor load, this will lead to variations in the specific energy and hence in pulp quality. Also, variations in the chip density reflect changes in the nature of the wood which may directly affect the pulp quality [61].

**Refiner plate wear**

Refiner operation may not have a stationary behavior partially due to refiner plate wearing. Plate wearing normally occurs gradually over a period of several hundred hours
due to erosion and corrosion, but it can also occur suddenly in the case of plate clashing [30]. For a given specific energy, new plates tend to produce the pulp with higher freeness than worn plates. When the plates become extremely worn, short-term load variations increase and it becomes increasingly difficult to fully load a refiner. Plate wearing cannot be avoided and often not even measured, but it can be compensated for by selecting some advanced control techniques.

2.2.5 Variable Interactions

Interactions between the variables are given in Figure 2.4. As can be seen, the final pulp quality is determined by the refining operating conditions and the raw material quality. Since variations in raw material quality cannot be controlled and often cannot even be...
measured, desired pulp quality can only be achieved by setting correct operating conditions. Changes in wood quality and operating conditions can be compensated through adjusting manipulated variables.

2.3 Modelling a TMP Process

The purpose of this section was to develop a steady state relationship between the variables described previously. The relationship developed not only gives some insight into process interactions and nonlinearity but also provides a valuable tool for process quantitative analysis. The variables appearing in the following derivation can be found in List of Symbols in the beginning of the thesis.

2.3.1 Modelling Throughput

Throughput or production rate is defined as the oven-dry wood flow fed to a refiner. Wood chips delivered by the chip transfer screw are usually metered on a volumetric basis. If assuming that the chip transfer screw is fully loaded with chips, the volumetric flow rate $Q$ of the wood chips (dry wood + moisture) to the process may be expressed as a function of the chip velocity $v$ in the screw direction:

$$Q = A \cdot v$$

where $A$ is the screw cross sectional area available to the chip flow.

Generally, the chip velocity $v$ in the screw direction is proportional to the chip screw rotational speed $S_c$, i.e.

$$v = k \cdot S_c$$

where $k$ is a constant factor, determined by the size of the screw systems. If assuming that the chip bulk density $\rho_c$ of the incoming wood is known, then the wood chip flow
rate \( f_c \) can be expressed by:

\[
f_c = \rho_c \cdot Q \tag{2.3}
\]

If assuming that the solid content \( s \) of the incoming chips is known, the oven-dry (moisture free) chip mass flow rate or throughput \( F_c \) is:

\[
F_c = s \cdot f_c \\
= s \cdot \rho_c \cdot Q \tag{2.4}
\]

Substituting Equations 2.1 and 2.2 into Equation 2.4 gives:

\[
F_c = s \cdot \rho_c \cdot A \cdot k \cdot S_c \tag{2.5}
\]

The above equation shows that for given chip screw systems, throughput or production rate \( F_c \) is a function of chip bulk density \( \rho_c \), chip solid content \( s \) and chip screw speed \( S_c \).

Generally, chip bulk density and solid content are not constant. Chip bulk density varies with changing in wood density and chip size, whereas solid content changes with changing in wood species and age of the chips. In practice, the chip transfer screw speed is the only available way to alter throughput or production rate.

### 2.3.2 Modelling Consistency

The consistency at the refiner outlet is derived based on the conservation principle of two fundamental quantities: the total mass and the total energy.

**The 1st-stage refining consistency**

The mass and energy flows in the primary refiner are illustrated in Figure 2.5. The energy balance in the primary refiner is:
energy in = energy out + energy losses

\begin{align}
\text{energy in} &= e_c + e_d + e_{sw} + \eta P \\
\text{energy out} &= e_{s1} + e_{s2} + e_p \\
\text{energy losses} &= 0 \text{ (assuming negligible energy losses)}
\end{align}

where

\begin{align}
e_c &= F_c \cdot c_c \cdot T_c + f_{w1} \cdot c_w \cdot T_c \\
e_d &= F_d \cdot c_w \cdot T_d \\
e_{sw} &= F_{sw} \cdot c_w \cdot T_{sw} \\
e_{s1} &= v_{s1} \cdot L \\
e_{s2} &= v_{s2} \cdot L \\
e_p &= f_f \cdot c_f \cdot T_p + f_{w2} \cdot c_w \cdot T_p
\end{align}

where \( v_{s1} \) is the backward steam flow rate, \( v_{s2} \) is the forward steam flow rate, and \( L \) is the enthalpy of the steam. Note that the temperatures given in Equation 2.7 are the relative temperatures from a reference temperature where the enthalpy of the liquids is assumed to be zero [68].
The mass balances in the primary refiner are:

**water in = water out + water losses**

water in = \( f_{w1} + F_d \)

water out = \( f_{w2} + v_{s1} + v_{s2} \)

water losses = \(-F_{sw}\) (seal water and leakage)

\[ (2.8) \]

**fibre in = fibre out + fibre losses**

fibre in = \( F_c \)

fibre out = \( f_f \)

fibre losses = 0 (assuming 100% yield pulp)

and the water flow rate \( f_{w1} \) carried by the inlet chips is calculated by:

\[ f_{w1} = \frac{F_c}{s} - F_c \]  \[ (2.9) \]

It can generally be assumed that:

\[ c_c = c_f \]

\[ T_d = T_{sw} \]  \[ (2.10) \]

Substituting Equations 2.7, 2.8, 2.9 and 2.10 into Equation 2.6 gives:

\[ \frac{v_{s1} + v_{s2}}{\eta P - \left[ F_c c_c (T_p - T_c) + (F_c/s - F_c) c_w (T_p - T_c) + (F_{sw} + F_d) c_w (T_p - T_d) \right]} \quad (L - c_w T_p) \]  \[ (2.11) \]

Consistency is defined as the ratio of dry fibre flow to the dry fibre flow plus water flow. So the outlet consistency \( C \) of the primary refiner is expressed as:

\[ C = \frac{f_f}{f_f + f_{w2}} \]  \[ (2.12) \]

Substituting Equations 2.8, 2.9 and 2.10 into 2.12 gives:

\[ C = \frac{F_c}{\frac{F_c}{s} + F_d + F_{sw} - (v_{s1} + v_{s2})} \]  \[ (2.13) \]
The above equation shows that the outlet consistency is a function of the dry chip flow $F_c$, chip solid content $s$, dilution flow rate $F_d$, seal water rate $F_{sw}$ and total steam flow $v_{s1} + v_{s2}$. Substituting Equation 2.11 into Equation 2.13 yields:

$$C = \frac{F_cL}{\frac{E_s}{s}L + F_dL + F_{sw}L - \frac{\eta P - [F_c c_c(T_p - T_c) + (F_c/s - F_c)c_w(T_p - T_c) + (F_{sw} + F_d)c_w(T_c - T_d)]}{1 - \frac{c_w T_p}{L}}}$$

(2.14)

In general

$$\frac{c_w T_p}{L} \ll 1$$

(2.15)

Equation 2.14 can be then simplified as:

$$C \approx \frac{F_cL}{\frac{E_s}{s}L + F_dL + F_{sw}L - \eta P}$$

(2.16)

Similarly, the inlet consistency $c_i$ is defined as the ratio of the dry wood flow rate to the total mass flow at inlet of a refiner:

$$c_i = \frac{F_c}{\frac{E_s}{s} + F_d + F_{sw}}$$

(2.17)

**The 2nd-stage refining consistency**

The consistencies in the 2nd stage refiner are derived in a similar fashion as were the consistencies in the 1st stage. The outlet consistency in the 2nd-stage refiner is given as:

$$C = \frac{F_cL}{\frac{F_c}{1-m}L + F_dL + F_{sw}L - \frac{\eta P - [F_c c_c(T_p - T_i) + \frac{(E_c - F_c)c_w(T_p - T_i)}{1-m} + (F_{sw} + F_d)c_w(T_p - T_d)]}{1 - \frac{c_w T_p}{L}}}$$

(2.18)

where $T_i$ (°C) is the temperature of the inlet pulp and $m$ (%) is the moisture content of the inlet pulp. If considering the conditions given in Equation 2.15, the above equation can be further simplified as:

$$C \approx \frac{F_cL}{\frac{F_c}{1-m}L + F_dL + F_{sw}L - \eta P}$$

(2.19)
The inlet consistency in the 2nd stage is calculated as:

\[ c_i = \frac{F_c}{\frac{F_{e}}{1-m} + F_d + F_{sw}} \]  \hspace{1cm} (2.20)

\textbf{Note:}

- The consistency equations derived above may be used to predict inlet or outlet consistency for each stage. If a consistency sensor is available and if significant errors between the measurements and model estimations exist, the errors should be taken into account in sensor calibration or consistency model improvement.

- The relationship between the inlet and outlet consistencies, e.g. in the 1st-stage, can be easily derived by substituting Equation 2.17 into Equation 2.16 as:

\[ C = \frac{L}{C_i / \eta P / F_c} \]  \hspace{1cm} (2.21)

If consistency measurements are available, the above equation can be used to back calculate the production rate \( F_c \), i.e.

\[ F_c = \frac{\eta P}{L} \left( \frac{C}{C - c_i} \right) \]  \hspace{1cm} (2.22)

Static modelling of the outlet consistency can also be found in [69, 70], but their models were derived with more complexity or unknown parameters. The consistency model given in [16] was derived based on the radial velocity of the pulp in the refining zone from a calculation of the forces that govern the flow of the pulp. However, simulations of the consistency model (Equation 2.16) based on mass-energy balances predict the same results as that given in [16].

\subsection*{2.3.3 Modelling Energy Input}

The degree of refining, a critical factor to the final product, is strongly influenced by the energy input to the pulp flow in wood chip refining.
Specific energy \((E)\) is defined as the ratio of the net power or the motor load \(P\) to the throughput or production rate \(F_c\):

\[
E = \frac{P}{F_c} \quad (2.23)
\]

The above equation shows that the specific energy \(E\) is proportional to the motor load \(P\) and inversely proportional to the throughput \(F_c\). Therefore, the motor load and/or the throughput can be adjusted according to the desired specific energy.

Motor load \((P)\) is defined as the electrical energy applied to a refiner. In a pressurized refiner, the total axial thrust force \(F_{total}\) applied on refiner plates is balanced by the steam reaction force \(F_s\) acted on the plates and the wood reaction force \(F_w\) exerted on the plates [71, 3, 72], i.e.

\[
F_{total} = F_w + F_s \quad (2.24)
\]

The experimental study by Miles and May shows that the motor load \(P\) for each refiner linearly depends on the wood reaction force \(F_w\) which is proportional to the closing pressure \(P_c\) of the hydraulic loading system. So we have:

\[
P = \alpha \cdot F_w \quad (2.25)
\]

and

\[
F_w = \beta \cdot P_c \quad (2.26)
\]

where \(\alpha\) and \(\beta\) are the coefficients. \(\alpha\) may be determined by the refiner size, disc rotational speed, refining zone consistency etc., whereas \(\beta\) is mainly determined by the size of the hydraulic loading system [71, 3, 72]. Combining the above two equations gives:

\[
P = \alpha \cdot \beta \cdot P_c \quad (2.27)
\]

Many experimental studies show that the motor load is also proportional to the throughput, i.e.

\[
P = k \cdot F_c \quad (2.28)
\]
This is because changes in the throughput will vary the wood reaction force on the plates and thereby alter the motor load.

Due to the complex relationship, little work has been done towards deriving the static quantitative relationship between the throughput, closing pressure and motor load. To start with a simple model, this relationship may be considered as:

\[ P = k P_c F_c \]  \hspace{1cm} (2.29)

For control design purposes, the above model can also be linearized about some operating point and modified as:

\[ P = k_1 P_c + k_2 F_c \]  \hspace{1cm} (2.30)

where \( P \) is the motor load deviation from its operating point, \( P_c \) and \( F_c \) are the deviations from their corresponding steady-state values, and \( k_1 \) and \( k_2 \) are the coefficients which will be determined by the operating conditions, plate conditions, refiner design and raw material quality.

**Residence time** (\( \tau \)) is defined as the time a pulp takes to pass through the refining zone. A theoretical model of the residence time has been derived by Miles and May [16] through calculation of the radial velocity of pulp in a high-consistency chip refiner. Through this model, the residence time is linked to a number of variables in terms of major operating variables and refiner design parameters as follows:

\[ \tau = \frac{\mu r}{\mu_t} \frac{a E c_i L_s}{w^3 (L_s (r_2^2 - r_1^2) + c_i E r_1^2)} \left[ \ln \frac{r_2}{r_1} - \frac{1}{2} \ln \left( \frac{L_s - c_i E}{L_s} \right) \right] \]  \hspace{1cm} (2.31)
where

\[ \begin{align*} 
  w & : \text{disc rotational speed} \\
  c_i & : \text{inlet consistency} \\
  r_1, r_2 & : \text{the inlet and outlet radii of the refining zone} \\
  \mu_r, \mu_t & : \text{radial and tangential friction coefficients between pulp and discs} \\
  L_s & : \text{latent heat of steam} \\
  a & : 4 \text{ for a single-disc refiner, } 2 \text{ for a double-disc refiner}
\end{align*} \]

As can be seen from the above equation, the residence time \( \tau \) is a function of refining consistency, disc rotational speed, specific energy and refiner design. Although the effect of rotational speed is far more significant, refining consistency, for a given specific energy, is the most readily available means of altering residence time in industrial operations.

**Refining intensity** \( (I) \) is defined as the average specific energy per refiner bar impact [16]:

\[ I = \frac{E}{N} \]  \hspace{1cm} (2.32)

where \( N \) is the number of refiner bar impacts received by unit pulp and is calculated by [55]:

\[ N = n \ h \ \frac{(r_1 + r_2)}{2} \ \tau \] \hspace{1cm} (2.33)

where

\[ \begin{align*} 
  n & : \text{the number of bars per unit length of arc} \\
  h & : 1 \text{ for a single-disc refiner, } 2 \text{ for a double-disc refiner.}
\end{align*} \]

It can be seen from Equation 2.32 that for a given specific energy \( E \), the number of bar impacts \( N \) governs the average amount of energy transferred at each impact, i.e. the refining intensity \( I \). The concept of refining intensity implies that refining will depend not only on the specific energy applied but also on the intensity of its application. Equations 2.32 and 2.33 show that for a given specific energy appropriate manipulation of the
residence time $\tau$ by adjusting inlet consistency will provide a desired combination of $N$ and $I$.

**Specific refining power** ($\dot{e}$) is defined as the rate of energy delivered per unit mass of pulp [6]:

$$\dot{e} = \frac{E}{\tau} \tag{2.34}$$

Equations 2.32 and 2.34 can be combined as a ratio:

$$\frac{\dot{e}}{I} = \frac{N}{\tau} \tag{2.35}$$

or

$$\dot{e} = n \cdot h \cdot w \frac{(r_1 + r_2)}{2} \cdot I \tag{2.36}$$

As can be seen from the above equation, the specific refining power is another measurement of the severity or intensity of refining for a given specific energy.

**Energy input over two stages.** The quality of final product is dependent not only on the energy input at each stage but also on total energy split and refining intensity distribution over two stages. The total motor load, specific energy, residence time and number of bar impacts can be expressed as:

$$P_{total} = P_1 + P_2$$
$$E_{total} = E_1 + E_2$$
$$\tau_{total} = \tau_1 + \tau_2$$
$$N_{total} = N_1 + N_2 \tag{2.37}$$

where the subscripts 1 and 2 refer to the first and the second stage respectively.

The total refining intensity $I_{total}$ and total specific refining power $\dot{e}_{total}$ are calculated by:

$$I_{total} = \frac{E_{total}}{N_{total}} \tag{2.38}$$
$$\dot{e}_{total} = \frac{E_{total}}{\tau_{total}} \tag{2.39}$$
2.3.4 Modelling Pulp Properties

Recent experimental work by Qian and Tessier [70] has shown that pulp properties such as freeness \(CSF(ml)\), long fibre content \(LF(\%)\) and shive content \(SC(\%)\) are mainly dependent on the specific energy \(E\), refining intensity \(I\) and specific refining power \(\dot{e}\). The empirical relationships between these variables are expressed as [70]:

\[
CSF = (CSF_0 - k_1(E - E_0))10^{k_2(I - I_0)} \\
LF = LF_0 + k_3(\dot{e} - \dot{e}_0) + k_4(E - E_0) \\
SC = SC_010^{k_5(E - E_0) + k_6(\dot{e} - \dot{e}_0)}
\]

where \(CSF_0\), \(LF_0\) and \(SC_0\) are the initial values of pulp properties. \(E_0\), \(I_0\) and \(\dot{e}_0\) are the initial conditions of operating variables. All the coefficients \(k_i\) \((i = 1 \cdots 6)\) are dependent on the type of refiner and can be determined or updated if on-line measurements of pulp properties and operating variables are available.

As can be seen from Equation 2.36, the specific refining power \(\dot{e}\) can be determined from the refining intensity \(I\). Therefore, the pulp quality in Equation 2.40 is mainly a function of the specific energy \(E\) and the refining intensity \(I\).

2.4 Simulation Results and Discussions

The objective of simulations was trying to demonstrate process interactions, nonlinearity and input-output quantitative relationship. Simulations were performed on the mathematical relationship illustrated in Figure 2.6, which was proposed in Section 2.3. In order to approximate features of an industrial plant, simulations were performed under the operating conditions of an industrial CTMP production line. Associated operating conditions and refiner design parameters are given in Tables\(^2\) 2.1 and 2.5.

\(^2\)see references [73] and [70]
Figure 2.6: Model structure of the 1st-stage wood-chip refining

2.4.1 Simulating the Primary Refiner

Effects of hydraulic closing pressure and chip transfer screw speed

Figure 2.7 shows for a given dilution flow rate \( F_d = 4.7kg/sec \) the effects of the first stage closing pressure \( P_c \) and chip transfer screw speed \( S_c \). The dashed line indicates \( S_c = 25 \text{ rpm} \ (F_c = 271 \text{ t/d}) \), the dashdot line indicates \( S_c = 27.5 \text{ rpm} \ (F_c = 298 \text{ t/d}) \) and the solid line indicates \( S_c = 30 \text{ rpm} \ (F_c = 325 \text{ t/d}) \).

In practice, an increase in the closing pressure for a given throughput will increase the motor load and hence the specific energy. This has been indicated in Figure 2.7. The reason is that increased closing pressure increases the total axial thrust force applied to the refiner plates and, as a result, increases the motor load which is proportional to the total thrust. The outlet consistency (Figure 2.7) increases with an increase in the closing pressure because a higher motor load resulting from increased closing pressure brings about a higher temperature in the refining zone and, therefore, more steam and less wet mass in the refining zone.
Table 2.1: Primary Refining Conditions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Typical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throughput, $F_c$</td>
<td>325 a.d.t/d</td>
</tr>
<tr>
<td>Specific Energy, $E$</td>
<td>3.4 GJ/t</td>
</tr>
<tr>
<td>Motor Load, $P$</td>
<td>12.5 MW</td>
</tr>
<tr>
<td>Dilution Water, $F_d$</td>
<td>280 l/min</td>
</tr>
<tr>
<td>Seal Water, $F_{sw}$</td>
<td>75 l/min</td>
</tr>
<tr>
<td>Disk Rotation Speed, $w$</td>
<td>1800 r/min</td>
</tr>
<tr>
<td>Chip Screw Speed, $S_c$</td>
<td>30 r/min</td>
</tr>
<tr>
<td>Chip Moisture Content</td>
<td>45.6±3.2 %</td>
</tr>
<tr>
<td>Chip Solid Content, $s$</td>
<td>54.4 %</td>
</tr>
<tr>
<td>Chip Bulk Density, $\rho_c$</td>
<td>98.9±5.8 kg/m$^3$</td>
</tr>
<tr>
<td>Closing Pressure, $P_c$</td>
<td>1100 psi</td>
</tr>
<tr>
<td>Refining Pressure</td>
<td>290 psi</td>
</tr>
<tr>
<td>Inlet Radii of Refining Zone, $r_1$</td>
<td>51.4 cm</td>
</tr>
<tr>
<td>Outlet Radii of Refining Zone, $r_2$</td>
<td>76.2 cm</td>
</tr>
<tr>
<td>Bars per Unit Length of Arc, $n$</td>
<td>100</td>
</tr>
<tr>
<td>Single Disc Refiner</td>
<td>$a=4$, $h=1$</td>
</tr>
</tbody>
</table>

As can be seen from Figure 2.7, an increase in the screw speed (or the throughput) increases the motor load and the outlet consistency. This is because increased throughput brings about a higher wood reaction force acted on the plates and hence a higher motor load. The higher the motor load is, the higher the temperature in the refining zone is, and as a result the less the wet mass and the higher the consistency are. As the specific energy is proportional to the motor load and inversely proportional to the throughput, the direction of changes in the specific energy is dependent on both the motor load and throughput. At a given closing pressure, the specific energy increases is due to a larger increase in the motor load, while the specific energy decreases is due to a larger increase in the throughput (see Figure 2.7(b)).

Table 2.2 summarizes the simulation results given in Figure 2.7. For a given throughput = 325 t/d, 40% increase in the closing pressure (from 1000 to 1400 psi) results in about 63% increase in the load (Figure 2.7(a)), 64% increase in the specific energy (Figure 2.7(b))
and 36% increase in the outlet consistency. Table 2.2 also shows that for a given closing pressure=1200 psi, 20% increase in the throughput (from 271 to 325 t/d) brings about 25% increase in the motor load, 5% increase in the specific energy, and 17% increase in the consistency. The simulation results show that the motor load seems more sensitive to the closing pressure changes than to the throughput.

<table>
<thead>
<tr>
<th>Table 2.2: First Stage Refining at $F_d = 4.7kg/sec$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closing Pres.=1000-1400psi △ Closing Pres.=40% (throughput=325t/d)</td>
</tr>
<tr>
<td>Motor Load (MW)</td>
</tr>
<tr>
<td>9.5 to 15.5</td>
</tr>
<tr>
<td>+63 %</td>
</tr>
</tbody>
</table>

Effects of dilution water and chip transfer screw speed

Figure 2.8 shows for a given specific energy ($E=3.4$ GJ/t) the effects of the dilution flow rate $F_d$ and the chip screw speed $S_c$. As can be seen, the inlet consistency decreases with increases in the dilution flow rate. For a given specific energy, the outlet consistency decreases with increases in the dilution flow rate for a given throughput. The explanation is that refining under constant specific energy and a given throughput determines a constant load applied. In this case, an increase in the dilution flow rate determines more wet mass in the refining zone and as a result decreases the outlet consistency.

Table 2.3 summarizes the simulation results given in Figure 2.8. For a given specific energy $E=3.4$ GJ/t, increasing dilution flow by 19% (from 4.2 kg/sec to 5.0 kg/sec) results in a 5.0% drop in the inlet consistency and a 7.8% drop in the outlet consistency. For a given dilution flow rate $F_d = 4.6$ kg/sec, increasing the chip screw speed by 20% (from 25 r/min to 30 r/min) brings about a 9.3% increase in the inlet consistency and a 16.5% increase in the outlet consistency. The simulation results indicate that the consistencies
Figure 2.7: The 1st stage refining at $F_d=4.7$ kg/sec: effects of closing pressure $P_c$ and chip screw speed $S_c$ (---: $S_c=25$ rpm ($F_c=271$ t/d), - - - : $S_c=27.5$ rpm ($F_c=298$ t/d) and ---: $S_c=30$ ($F_c=325$ t/d))
are more sensitive to the chip feed or the throughput.

Figure 2.8: The 1st stage refining at $E = 3.4$ GJ/t: effects of dilution flow rate $F_d$ and chip screw speed $S_c$ (- - : $S_c = 25$ rpm ($F_c = 271$ t/d), - - - : $S_c = 27.5$ rpm ($F_c = 298$ t/d) and --- : $S_c = 30$ ($F_c = 325$ t/d)).

Table 2.3: First Stage Refining at $E = 3.4$ GJ/t

<table>
<thead>
<tr>
<th>$F_d$ = 4.2 - 5.0 kg/sec</th>
<th>$S_c$ = 25 - 30 r/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta F_d = 19%$</td>
<td>$\Delta S_c = 20%$</td>
</tr>
<tr>
<td>$(S_c = 30$ r/min)</td>
<td>$(F_d = 4.6$ kg/sec)</td>
</tr>
<tr>
<td>$c_i$ (%) $C$</td>
<td>$c_i$ (%) $C$</td>
</tr>
<tr>
<td>30.1 to 28.6 51 to 47</td>
<td>27 to 29.5 42.5 to 49.5</td>
</tr>
<tr>
<td>-5 % -7.8 %</td>
<td>+9.3 +16.5 %</td>
</tr>
</tbody>
</table>

Effects of rotational speed and inlet consistency

Figure 2.9 shows for a given specific energy $E = 3.4$ GJ/t the influences of the rotational speed $w$ and inlet consistency $c_i$ on operating variables and pulp properties. The solid line indicates $c_i = 20\%$, dashdot line $c_i = 23.5\%$ and dashed line $c_i = 27\%$. 
As can be seen from Figure 2.9, an increase in the rotational speed results in decreases in the residence time and the number of bar impacts, and increases in the refining intensity and the specific refining power. The explanation is that increased rotational speed increases the centrifugal force on the pulp flow in the refining zone and thus increases the radius velocity of the pulp. As a result, the residence time decreases and thus the number of refiner bar impacts received by unit pulp decreases. For a given specific energy, the specific energy per impact, i.e. the refining intensity increases due to a decrease in the number of bar impacts. The specific refining power increases as a result from a decrease in the residence time. Increased refining intensity will bring about more fibre cutting. As a result, there are less long fibres and shives but more fines, resulting in a decrease in the pulp drainage or freeness.

For a given specific energy \( E = 3.4GJ/t \), an increase in the inlet consistency increases the residence time and thus the number of bar impacts. As a result, the refining intensity and the specific refining power decrease. This is because increased inlet consistency reduces the wet mass in the refining zone, resulting in a decrease in the centrifugal force on the pulp flow in the refining zone and hence increases in the residence time and the number of bar impacts. For a given specific energy, increased number of bar impacts brings about a decrease in the refining intensity, whereas an increase in the residence time decreases the specific refining power. Decreased refining intensity results in less fibre cut and thus more long fibres but less fines. Therefore, the freeness, long fibre content and shive content are increased.

Table 2.4 summarizes the simulation results given in Figure 2.9. At a given inlet consistency \( c_i = 27\% \), a 29% increase in the rotational speed (from 1400 rpm to 1800 rpm) brought about a 58% drop in the residence time \( \tau \) (from 0.6 sec to 0.25 sec). As a result, the number of impacts \( N \) was reduced by about 44% (from 5500 to 3100) and the refining intensity \( I \) increased by about 67% (from \( 0.6 \times 10^{-3} \) MJ/kg to \( 1.0 \times 10^{-3} \)).
Figure 2.9: The 1st stage refining at E=3.4 GJ/t: effects of rotational speed \( w \) and inlet consistency \( c_i \) (---: \( c_i = 20\% \), -.-.: \( c_i = 23.5\% \) and --: \( c_i = 27\% \))
MJ/kg). The changes in the specific refining power were even more pronounced, which was a 127% increase from 5.5 MJ/kg/sec to 12.5 MJ/kg/sec. As a result of a 28.6% increase in the rotational speed, freeness reduced by 38%, long fibre content reduced by 23% and shive content decreased by 50%.

Table 2.4 shows that at a given rotational speed, e.g. $w = 1800$, increasing the inlet consistency by about 35% (from 20% to 27%) results in a 25% increase in the residence time $\tau$, 44% increase in the number of impacts $N$ and 33% decrease in the refining intensity. A 35% increase in the inlet consistency also brings about a 67% increase in the freeness CSF, 25% increase in the long fibre content LF and 100% increase in the shive content SC. The simulation results indicate that for a given specific energy refining conditions and pulp properties are strongly related to the rotational speed and refining consistency. The effect of rotational speed seems to be more apparent. If there were a technique available for on-line adjusting the rotational speed, there would be a more sensitive way for the control of refining process. However, the consistency at present is the only available way to alter pulp properties for a given specific energy.

<table>
<thead>
<tr>
<th>$w = 1400 - 1800$ r/min</th>
<th>$\Delta w = 29%$</th>
<th>$c_i = 27%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$ (sec)</td>
<td>$N$</td>
<td>$I$ ($\times 10^{-3}$ MJ/kg)</td>
</tr>
<tr>
<td>0.6 to 0.25</td>
<td>5500 to 3100</td>
<td>0.6 to 1.0</td>
</tr>
<tr>
<td>-58%</td>
<td>-44%</td>
<td>+67%</td>
</tr>
<tr>
<td>$c_i = 20 - 27%$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta c_i = 35%$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w = 1800$ r/min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$ (sec)</td>
<td>$N$</td>
<td>$I$ ($\times 10^{-3}$ MJ/kg)</td>
</tr>
<tr>
<td>0.2 to 0.25</td>
<td>2250 to 3250</td>
<td>1.5 to 1.0</td>
</tr>
<tr>
<td>+25%</td>
<td>+44%</td>
<td>-33%</td>
</tr>
</tbody>
</table>
2.4.2 Simulating the Secondary Refiner

Effect of rotational speed and inlet consistency

Figures 2.10 and 2.11 show for a given secondary specific energy $E_2 = 2.5 GJ/t$, the effects of the 2nd-stage rotational speed $w$ and the inlet consistency $c_i$. The solid line indicates $c_i = 26\%$, dashdot line $c_i = 30.5\%$ and dashed line $c_i = 35\%$. In all simulations, the specific energy and inlet consistency in the 1st stage were set to $E_1 = 3.4 GJ/t$ and $c_1 = 27\%$ respectively which produced the residence time $\tau_1 = 0.2753\ sec$ and the number of impacts $N_1 = 3.3 \times 10^3$.

<table>
<thead>
<tr>
<th>Table 2.5: Secondary Refining Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td>Specific Energy, $E$</td>
</tr>
<tr>
<td>Motor Load, $P$</td>
</tr>
<tr>
<td>Dilution Rate, $F_d$</td>
</tr>
<tr>
<td>Seal Water, $F_{sw}$</td>
</tr>
<tr>
<td>Disk Rotation Speed, $w$</td>
</tr>
<tr>
<td>Closing Pressure, $P_c$</td>
</tr>
<tr>
<td>Inlet Radii of Refining Zone, $r_1$</td>
</tr>
<tr>
<td>Outlet Radii of Refining Zone, $r_2$</td>
</tr>
<tr>
<td>Bars per Unit Length of Arc, $n$</td>
</tr>
<tr>
<td>Single Disc Refiner</td>
</tr>
</tbody>
</table>

Figure 2.10 shows that an increase in the rotational speed (or inlet consistency) results in decreases (or increases) in the residence time and the number of bar impacts, but increases (or decreases) in the refining intensity and the specific refining power. The explanation is the same as that given for Figure 2.9. Decreases in the residence time $\tau_2$ and number of bar impacts $N_2$ in the 2nd-stage (Figure 2.10) resulted in decreases in the total residence time $\tau_t = \tau_1 + \tau_2$ and total number of bar impacts $N_t = N_1 + N_2$. As a result, the refining intensity $I_t$ and specific refining power $\dot{e}_t$ over two stages increased. Increased total refining intensity (or specific refining power) brought up more fibre cut...
and hence less long fibres and shives but more fines. As a result, the freeness, long fibre content and shive content decreased.

Table 2.6 summarizes the simulation results given in Figure 2.10. For a given inlet consistency $c_i = 30.5\%$, a 28.6\% increase in the rotational speed $w_2$ (from 1400 rpm to 1800 rpm) decreased the residence time $\tau_2$ by about 53\% (from 0.475 sec to 0.225 sec) and the number of impacts by about 36\% (from 4250 to 2700). As a result, the refining intensity $I_2$ increased by about 69.6\% (from $5.6 \times 10^{-4}$ MJ/kg to $9.5 \times 10^{-4}$ MJ/kg). The increase in the specific refining power $\dot{e}_2$ was even more pronounced, which is about 105\% from 5.5 MJ/kg/s to 11.3 MJ/kg/s. The effect of the inlet consistency is smaller compared to that of the rotational speed. Table 2.6 shows that at a given rotational speed, e.g. $w = 1800$ r/min, a 35\% increase in the inlet consistency (from 26\% to 35\%) resulted in a 39\% increase in the residence time, a 41\% increase in the number of impacts, a 29\% decrease in the refining intensity and a 28.7\% decrease in the specific refining power, respectively. Table 2.7 shows that a 28.6\% increase in the secondary rotational speed resulted in a 33\% decrease in the total residence time $\tau_t$, a 24\% decrease in the total number of impacts $N_t$ and a 28\% increase in the total refining intensity $I_t$. In terms of pulp properties, a 28.6\% increase in the secondary rotational speed decreased the 2nd stage freeness by 50\%, the long fibre content by 18\% and the shive content by 44\%. At a given rotational speed $w_2 = 1800$ r/min, Table 2.7 shows that a 35\% increase in the inlet consistency (from 26 to 35\%) on the 2nd stage brought about a 22\% increase in $\tau_t$, an 18\% increase in $N_t$ and 17.4\% decrease in $I_t$. As a result, the freeness, the long fibre content and the shive content on the secondary stage increased by 21\%, 11\%, and 31\% respectively.
Figure 2.10: The 2nd stage refining at $E_2 = 2.5$ GJ/t: effects of the rotational speed $w$ and the inlet consistency $c_i$ (---: $c_i = 26\%$, -.-.: $c_i = 30.5\%$ and - - : $c_i = 35\%$)
Figure 2.11: Two-stage refining at $E_2 = 2.5$ GJ/t: effects of the 2nd-stage rotational speed $w$ and the inlet consistency $c_i$ (— $c_i = 26\%$, -.-: $c_i = 30.5\%$, and - - $c_i = 35\%$)
Chapter 2. Process Variables, Modelling and Simulations

Table 2.6: Effects of Secondary Stage Refining at $E=2.5$ GJ/t

$w_2 = 1400 - 1800$ r/min

$\Delta w_2 = 29\%$

($c_i = 30.5\%$)

<table>
<thead>
<tr>
<th>$\tau_2$ (sec)</th>
<th>$N_2$</th>
<th>$I_2$ ($\times 10^{-4}$ MJ/kg)</th>
<th>$\dot{e}_2$ (MJ/kg/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.475 to 0.225</td>
<td>4250 to 2700</td>
<td>5.6 to 9.5</td>
<td>5.5 to 11.3</td>
</tr>
<tr>
<td>-53%</td>
<td>-36%</td>
<td>+69.6%</td>
<td>+105%</td>
</tr>
</tbody>
</table>

$c_i = 26 - 35\%$

$\Delta c_i = 35\%$

($w_2 = 1800$ r/min)

<table>
<thead>
<tr>
<th>$\tau_2$ (sec)</th>
<th>$N_2$</th>
<th>$I_2$ ($\times 10^{-4}$ MJ/kg)</th>
<th>$\dot{e}_2$ (MJ/kg/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.18 to 0.25</td>
<td>2200 to 3100</td>
<td>11.3 to 8.0</td>
<td>13.6 to 9.7</td>
</tr>
<tr>
<td>+39%</td>
<td>+41%</td>
<td>-29%</td>
<td>-28.7%</td>
</tr>
</tbody>
</table>

Table 2.7: Effects of Secondary Stage Refining on Pulp Properties, $E=2.5$ GJ/t

$w_2 = 1400 - 1800$ r/min

$\Delta w_2 = 29\%$

($c_i = 30.5\%$)

<table>
<thead>
<tr>
<th>$\tau$ (sec)</th>
<th>$N_t$</th>
<th>$I_t$ ($\times 10^{-4}$ MJ/kg)</th>
<th>CSF (ml)</th>
<th>LF (%)</th>
<th>SC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75 to 0.5</td>
<td>7900 to 6000</td>
<td>7.8 to 10</td>
<td>200 to 100</td>
<td>45 to 37</td>
<td>0.27 to 0.15</td>
</tr>
<tr>
<td>-33%</td>
<td>-24%</td>
<td>+28%</td>
<td>-50%</td>
<td>-18%</td>
<td>-44%</td>
</tr>
</tbody>
</table>

$c_i = 26 - 35\%$

$\Delta c_i = 35\%$

($w_2 = 1800$ r/min)

<table>
<thead>
<tr>
<th>$\tau$ (sec)</th>
<th>$N_t$</th>
<th>$I_t$ ($\times 10^{-4}$ MJ/kg)</th>
<th>CSF (ml)</th>
<th>LF (%)</th>
<th>SC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.45 to 0.55</td>
<td>5500 to 6500</td>
<td>10.9 to 9.0</td>
<td>140 to 170</td>
<td>35 to 39</td>
<td>0.13 to 0.17</td>
</tr>
<tr>
<td>+22%</td>
<td>+18%</td>
<td>-17.4%</td>
<td>+21%</td>
<td>+11%</td>
<td>+31%</td>
</tr>
</tbody>
</table>
2.5 Summary and Conclusions

The purpose of this chapter was to obtain the quantitative relationships between process inputs and outputs. A set of equations was derived using the mechanistic method, through which major manipulating variables are linked to major output variables. As can be seen from process modelling, for given raw material and refiner design the pulp quality is mainly determined by specific energy and the rate of transferred energy, i.e., refining intensity (or specific refining power). Specific energy is a function of refiner motor load and throughput, while refining intensity is a function of specific energy, disc rotational speed and refining consistency. Although simulations showed that the effect of the rotational speed is the most apparent one, for a given specific energy refining consistency appeared to be the only readily available means of adjusting refining intensity in an industrial operation. Simulations also demonstrated the quantitative relationships between the variables and process nonlinearity.
Chapter 3

Industrial TMP Process Identification

3.1 Introduction

The model developed in the previous chapter provides a valuable tool for qualitative analysis and a better understanding of the process. For control purposes, process dynamics and disturbances should be considered in the model development. However, it is not easy to model the dynamics of the process through mechanistic modelling and it is even impossible to characterize the disturbances by an analytical (or mechanistic) function [74]. A dynamic model which possesses maximum simplicity and a minimum number of parameters consonant with representational adequacy is usually achieved from process identification, i.e., a model is derived experimentally by using input-output data collected from a real process (Figure 3.12).

One of the advantages in using process identification or the empirical method is that only quantitative knowledge of the process mechanism is required to find rough estimates of process input-output relationships. Secondly, using process identification, dynamic modelling can be realized more easily in less time than using the mechanistic analysis where the development of complex differential equations may be required. Lastly, due to the impossibility of modelling disturbances by the mechanistic method, process identification becomes a successful tool in characterizing disturbances.

Conventional identification methods are based on the choice of special inputs to a process, such as step, pulse and sinusoidal inputs. These methods have been useful
when a process is affected by a small disturbance. However, they have not always been successful [75] and usually provide poor results on processes characterized by long time constants or excessive measurement noises. This is because large process input manipulations are needed to generate the required information at process outputs, thus risking sustained deviations from product quality targets and unsafe operating conditions. The approach adopted in this study is time series analysis which refers to the analysis of data from random observations with time (Figure 3.12). The reasons for using time series techniques are capabilities of (i) identifying accurate models even when the input variations are small, (ii) modelling the effects of all unmeasured external disturbances and measurement noises, and (iii) providing a general methodology for identification of the multivariate processes.

The objective of industrial TMP process identification is to find a dynamic-stochastic model suitable for the control of a TMP plant. To achieve this goal, a number of mill trials have been performed at a B.C. pulp and paper mill. A large number of on-line operational data have been collected. With the help of identification techniques and time series analysis, intensive data analysis has been performed to interpret the data. Section 3.2 describes the choice of a model structure for process identification. Section 3.3 details the time series analysis techniques used for process identification. Section 3.4 presents the procedures for identifying, fitting, and checking models when simultaneous
pairs of observations \((u_1, y_1), (u_2, y_2), \cdots, (u_t, y_t)\) of input and output are available. Section 3.5 presents identification results and discussions of the results. The summary and conclusions are given in Section 3.6.

### 3.2 Process Model

Before performing process identification, it is always necessary to assume the type of model according to the prior knowledge of the process and the purpose of the model development. For control purposes, we assume that a stochastic process (simply referred to a process) can be described by a discrete-time dynamic-disturbance model as follows for a SISO case:

\[
y_t = G_p(z^{-1})u_{t-k-1} + N_t
\]

(3.41)

where \(y_t\) is the output deviation from its setpoint at time \(t\) and \(u_t\) is the input deviation from its corresponding steady-state value. The parameter \(k(\geq 0)\) is the number of sampling intervals of pure process delay. One additional period of delay is introduced by the sample and hold operation of the computer. \(G_p(z^{-1})\) presents the input-output dynamics of the process, such as the process gain and time constant. \(N_t\) characterizes all unmeasured external disturbances and measurement noise as observed at the output. Box and Jenkins [75] suggest that the transfer function-stochastic model above be parameterized as:

\[
y_t = \frac{B(z^{-1})}{A(z^{-1})} u_{t-k-1} + \frac{C(z^{-1})}{D(z^{-1})\Delta^d} e_t
\]

(3.42)

As can be seen, the disturbance term \(N_t\) in Equation 3.41 is modelled by passing an independent, normally distributed (white) noise sequence \(e_t\) with variance \(\sigma_e^2\) through a linear dynamic transfer function \(\frac{C(z^{-1})}{D(z^{-1})\Delta^d}\) \((\Delta = 1 - z^{-1})\). The advantages of using the Box-Jenkins model are that first, the allowance of \(d\) roots equal to unity \((d = 0\) or \(1\) usually) enables one to characterize the type of nonstationary drifting behavior that process
variables tend to exhibit when uncontrolled; secondly, because an integrator function is modelled in the disturbance model, it naturally leads to integral action in a model-based controller design. Finally, by using the Box-Jenkins model, the effects of both manipulated input and disturbance can be modelled separately. Terms $A(z^{-1}), B(z^{-1}), C(z^{-1})$ and $D(z^{-1})$ are the polynomials in the backward shift operator $z^{-1}$ (i.e., $z^{-1}y_t = y_{t-1}$) and they are described as follows:

$$
A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{na} z^{-na} \\
B(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{nb} z^{-nb} \\
C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_{nc} z^{-nc} \\
D(z^{-1}) = 1 + d_1 z^{-1} + d_2 z^{-2} + \cdots + d_{nd} z^{-nd}
$$

(3.43)

3.3 Analysis Tools

Once the data has been collected, a number of statistical analyses are performed, consisting of (i) the auto-correlations, (ii) the partial auto-correlations, and (iii) the cross-correlation functions. These are aimed primarily at identifying the model structure (Equation 3.42) from input-output data, and then performing model parameter estimation and model checking. The calculation of the auto-, partial auto- and cross-correlations are given in Appendix A, and associated properties can be found in the book [75] by Box and Jenkins.

3.4 Identification Procedures

Once the data has been collected, the steps involved in model identification are (i) identification of the model structure ($na, nb, nc, nd, k$), (ii) model parameter estimation, and (iii) identified model checking. The following presents the details of each step in process model identification.
3.4.1 Model Structure Identification

_Differencing Data to Obtain Stationarity_

The time series analysis tools given in Section 3.3 are based on the assumption of the stationarity of a process, which implies that the process has constant mean $\mu$ and constant variance $\sigma^2$. In practice, however, this is not always true. The nonstationarity of a process exists when time series auto- and cross-correlations do not die away, but rather are composed of many significant correlations even at high lags. Alternatively, the nonstationary behavior is suspected when a time series does not vary about a fixed mean value, but rather slowly meanders from one operation point to another. The stationarity of a process can be achieved by differencing the process data $(U_t, Y_t)$ pair $d$ times ($d$ is usually 1 or 2) until the stationary data $(u_t, y_t)$ pair is obtained, where $u_t = \Delta^d U_t$ and $y_t = \Delta^d Y_t$.

_Identifying the Structure of a Transfer Function_

The structure $(na, nb, k)$ (Equation 3.43) of a transfer function is obtained by estimating the impulse responses of the process and then comparing them with theoretical ones. Theoretical examples of impulse responses for some common processes can be found in [75].

In the case of open-loop process identification, if the input series $u_t$ is white (uncorrelated), then the process impulse response is proportional to the cross-correlation $r_{uy}(k)$ and can be obtained by multiplying the cross-correlation coefficients $r_{uy}(k)$ by $\sigma_y/\sigma_u$. However, if the input signal $u_t$ is not white, prewhitening the input has to be done before the estimation of impulse response. Prewhitening can be obtained by fitting the input $u_t$ to a time series model, i.e., $u'_t = \phi(z^{-1})\theta^{-1}(z^{-1}) u_t$ and transforming the correlated input series $u_t$ to the uncorrelated white series $u'_t$. The same transformation is applied to the output series $y_t$ to obtain $y'_t$. After prewhitening, the cross-correlation function between
the input $u'_t$ and correspondingly transformed output $y'_t$ is proportional to the impulse response.

Alternatively, the impulse response can be obtained without prewhitening by relating the input-output cross-covariance to the input auto-covariance. Suppose that a process can be adequately approximated by the following transfer function:

$$
y_t = v(0)u_t + v(1)u_{t-1} + v(2)u_{t-2} + \cdots
$$

$$= [v(0) + v(1)z^{-1} + v(2)z^{-2} + \cdots]u_t
$$

(3.44)

where the weights $v(0), v(1), v(2) \cdots$ in the above equation are called the impulse response. Multiplying the above equation by $u_{t-k}$ for $k \geq 0$ and taking expectations gives:

$$c_{uy}(k) = v(0)c_{uu}(k) + v(1)c_{uu}(k - 1) + \cdots
$$

$$k = 0, 1, 2, \cdots
$$

(3.45)

Supposing that the weights $v(i)$ are effectively zero beyond $k = K$, then the above equation with $k = 1, 2, 3 \cdots K$ can be written in a more compact form as:

$$C_{uu} v = c_{uy}
$$

(3.46)

or

$$v = C_{uu}^{-1} c_{uy}
$$

(3.47)

where

$$C_{uu} = \begin{bmatrix}
    c_{uu}(0) & c_{uu}(1) & \cdots & c_{uu}(K) \\
    c_{uu}(1) & c_{uu}(0) & \cdots & c_{uu}(K-1) \\
    \vdots & \vdots & \ddots & \vdots \\
    c_{uu}(K) & c_{uu}(K-1) & \cdots & c_{uu}(0)
\end{bmatrix}
$$

$$c_{uy} = \begin{bmatrix}
    c_{uy}(0) \\
    c_{uy}(1) \\
    \vdots \\
    c_{uy}(K)
\end{bmatrix}
$$

$$v = \begin{bmatrix}
    v(0) \\
    v(1) \\
    \vdots \\
    v(K)
\end{bmatrix}
$$

Identifying the Structure of a Disturbance Model
The idea of identifying the structure of a noise model (Equation 3.42) is the same as the idea of identifying the structure of an input-output transfer function. Once an acceptable input-output transfer function is obtained (see Section 3.4.3), the structure \((nc, nd)\) of the noise model can be obtained through calculating the residual auto- and partial auto-correlations and comparing them to the theoretical ones of known processes. Theoretical examples of the auto- and partial auto-correlations of some known processes can be found in [75].

3.4.2 Parameter Estimation

Once an acceptable structure \((na, nb, nc, nd, k)\) of a dynamic-noise model is obtained, estimation of the parameters of \(A(z^{-1}), B(z^{-1}), C(z^{-1})\) or \(D(z^{-1})\) polynomials in Equation 3.42 is carried out by minimizing the sum of prediction error as follows:

\[
\min \sum_{i=1}^{N} \hat{e}_t^2
\]

where

\[
\hat{e}_t = \frac{\hat{D}(z^{-1})}{\hat{C}(z^{-1})}[y_t - \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})}u_{t-k}]
\]

where \(\hat{e}_t\) is calculated at given \((y_t, u_{t-k})\) data pair and a set of initial parameters of \(A(z^{-1}), B(z^{-1}), C(z^{-1})\) and \(D(z^{-1})\) polynomials.

3.4.3 Model Checking

The tools used for checking the accuracy of a model are the residual auto-, partial auto- and cross-correlations. An adequate input-output model is obtained when the residual cross-correlations die away within a defined confidence interval. Similarly, a disturbance model is adequate when the residual auto- and partial auto-correlations die away within the confidence interval.
3.5 Industrial TMP Process Identification

The TMP process identification aimed at identifying the dynamics between inputs and outputs, including model order, time-delay, time constant and process gain. A multivariable relationship between inputs and outputs was obtained through identifying the relationship for each input-output pair by doing experiments separately. The following presents the details of the mill trials, on-line data and model identification results.

3.5.1 Process Description and Experiment Design

A number of mill trials were conducted on the primary refiner of a two-stage CTMP production line at a B.C. pulp and paper mill. Heated chips were fed from a vertical steaming vessel to a stream splitter conveyor by a pressurized plug screw discharger. From the stream splitter conveyor, the chips were metered by two load sensing conveyors into the drive and tail ends of the pressurized primary refiner (a twin-disc refiner). Dilution water was sent into the drive and tail ends of the primary refiner to obtain the desired consistency. From the primary refiner, pulp was blown into the pressurized primary cyclone and was then fed to the pressurized secondary refiner.

Because of pulp quality specifications, the process was operated under the conditions of minimizing the deviations of the process variables from their targets. This made the identification more difficult because of the lack of excitation for the process identification. Therefore, Pseudo Random Binary Sequences (PRBS) – a series of uncorrelated input changes, were added as input perturbation signals during identification. The advantage of using PRBS signals rather than a classical bump is that PRBS signals of much lower amplitude can be used, thereby minimizing the effects of the test on the process. Furthermore, because a PRBS signal is continually switching states, the process is excited during the entire time history rather than just at the start of the record as is the case
during traditional bump tests.

The amplitude of a PRBS perturbation signal is an important design variable and may be designed by placing hard limits on the input variations, i.e., $u_L \leq u_t \leq u_U$ in order to minimize process deviations from the desired targets while generating as much information on the process dynamics as possible. The upper and lower limits should be chosen based on how much variation is tolerable. The other design variable is the switching period $T$ of a PRBS signal. In order to have an input with a proper frequency content, the switching period $T$ should, in general, be some integer multiple of the sampling period $T_s$, i.e., $T = n \times T_s$ [76]. For the duration of our mill trials, the amplitudes of PRBS signals were limited to $\pm 0.8\%$ of each input and the switching period $T$ was chosen as two times of the sampling rate, i.e., $T = 2 T_s$. Key operating data was collected from Bailey Distributed Control Systems (DCS), where the sampling rate was one second, i.e., $T_s = 1$ sec.

### 3.5.2 Identification Results

The initial plan of process identification was to identify the dynamic relationship between inputs (closing pressure, chip screw speed and dilution water) and outputs (motor load and consistency). The simplest way for doing so is to identify the relationship between each input and output separately by doing the experiments separately, i.e., perturbing one input at a time while keeping other inputs constant.

**Closing Pressure (CP) to Motor Load (ML)**

Figure 3.13 shows the refiner motor load (ML) response to the changes in the hydraulic closing pressure (CP). Figure 3.14 shows the unit impulse response and its 99% confidence level, estimated using the motor load/closing pressure data from Figure 3.13. In practice, increasing CP increases ML. This has been indicated by the first few weights of the impulse responses. The second, third and forth impulse weights are significant, indicating
Table 3.8: ML/CP Parameter Estimates, Estimated Standard Deviations and FPE

<table>
<thead>
<tr>
<th>Model (Eqn.)</th>
<th>d</th>
<th>A(z&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>B(z&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>C(z&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>D(z&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>k</th>
<th>FPE std</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.50</td>
<td>0</td>
<td>1 -0.8678</td>
<td>0.0061</td>
<td>1 0.3941</td>
<td>1 -0.5096</td>
<td>1</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.1159</td>
<td>0.0026</td>
<td>0 0.0867</td>
<td>0 0.0831</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.51</td>
<td>1</td>
<td>1 -0.8852</td>
<td>0.0060</td>
<td>1 0.1473</td>
<td>1</td>
<td>1</td>
<td>0.1719</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.3141</td>
<td>0.0030</td>
<td>0 0.0711</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

that the process may have one or more complete sampling intervals of pure time delay (k ≥ 1). The remaining impulse weights suggest one of several alternatives for the orders of the transfer function polynomials A(z<sup>-1</sup>) and B(z<sup>-1</sup>). The two most likely alternatives are first order or second order. Note that the impulse response (Figure 3.14) also shows a periodic variation of 5 sec in the motor load. This variation might result from the periodic variation in the chip feed.

An iterative model-building procedure consisting of structure identification, fitting, and model checking stages identified the following dynamic-stochastic model:

\[
y_t = \frac{0.0061}{1 - 0.8678 z^{-1}} u_{t-2} + \frac{1 + 0.3941 z^{-1}}{1 - 0.5096 z^{-1}} e_t \tag{3.50}
\]

Table 3.8 shows the parameter estimates, their estimated standard deviations and Akaike’s FPE (Final Prediction Error) [77]. The above equation was further tested on a fresh data set that was not used in the model estimation. Associated residual tests in Figure 3.15 indicate that the above dynamic-stochastic model is adequate because the cross-correlations all fall in the 99% confidence intervals and the residuals are nearly white.

The input-output transfer function part, however, may be of limited use if it was used to design a minimum variance controller or a generalized predictive controller. The controller would not contain integral action because of the lack of nonstationarity in the disturbance model. If this were the intended use of the model then it would be necessary
to force an integrator into the disturbance model even though this may not be entirely supported by the data. Fitting the noise model with an integrator gave the following parameter estimates:

\[
y_t = \frac{0.006}{1 - 0.8852z^{-1}} u_{t-2} + \frac{1 + 0.1473z^{-1}}{\Delta} e_t
\]  

(3.51)

Table 3.8 shows the parameter estimates, estimated standard deviations and Akaike's FPE. The residual correlations calculated using a fresh data set are given in Figure 3.16, indicating that the above model is adequate. Figure 3.17 shows the fresh data set and the model prediction using the same data set.

![Figure 3.13: Motor Load (ML) response to Closing Pressure (CP)](image)
Chapter 3. Industrial TMP Process Identification

Figure 3.14: ML/CP impulse response estimate

Figure 3.15: ML/CP residual correlations (d=0)
Chapter 3. Industrial TMP Process Identification

Figure 3.16: ML/CP residual correlations (d=1)

Figure 3.17: CP/ML input, output and model prediction
Table 3.9: ML/TS Parameter Estimates, Estimated Standard Deviations and FPE

<table>
<thead>
<tr>
<th>Model (Eqn.)</th>
<th>d</th>
<th>$A(z^{-1})$ std</th>
<th>$B(z^{-1})$ std</th>
<th>$C(z^{-1})$ std</th>
<th>$D(z^{-1})$ std</th>
<th>k</th>
<th>FPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.52</td>
<td>0</td>
<td>1 -0.8186</td>
<td>0.0616</td>
<td>1 0.5900</td>
<td>1 -0.6159</td>
<td>8</td>
<td>0.1148</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 0.2527</td>
<td>0.0563</td>
<td>0 0.0643</td>
<td>0 0.0665</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.53</td>
<td>1</td>
<td>1 -0.8997</td>
<td>0.0548</td>
<td>1 0.5768</td>
<td>1</td>
<td>8</td>
<td>0.1425</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 0.4630</td>
<td>0.0672</td>
<td>0 0.0561</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Transfer Screw Speed (TS) to Motor Load (ML)**

Figure 3.18 shows the motor load (ML) response to the changes in the transfer screw speed (TS). To have an initial guess of the model structure, the unit impulse responses are estimated and given in Figure 3.19, using the raw data in Figure 3.18. In practice, increasing the chip feed speed increases the motor load. The first eight impulse weights in Figure 3.19 are insignificant, indicating that the process may have eight sampling period intervals of pure time delay ($k = 8$). This delay is attributed to the finite time required to deliver wood chips to the inlet of the primary refiner. The remaining impulse weights suggest one of several alternatives for the orders of the transfer function polynomials. The most likely alternatives are first order with fractional delay, or second order.

An iterative modelling procedure consisting of structure identification, model fitting, and model checking stages identified the following dynamic-stochastic model structure:

$$y_t = \frac{0.0616}{1 - 0.8186z^{-1}} u_{t-9} + \frac{1 + 0.59z^{-1}}{1 - 0.6159z^{-1}} e_t$$  (3.52)

Table 3.9 shows the parameter estimates, their estimated standard deviations and Akaike's FPE. A check on the above model was made by calculating the residual auto- and cross-correlations with the use of a fresh data set that was not used for the model estimation. The residual correlations in Figure 3.20 indicate that the identified the dynamic-stochastic model (Equation 3.52) is adequate.
To design a controller containing integral action, we forced an integrator into the disturbance model and obtained:

\[
y_t = \frac{0.0548}{1 - 0.8997z^{-1}} u_{t-9} + \frac{1 + 0.5768z^{-1}}{\Delta} e_t
\]

(3.53)

Table 3.9 shows the parameter estimates, estimated standard deviations and Akaike's FPE. A check on the above model was made by calculating the residual auto- and cross-correlations with the use of a fresh data set. The residual correlations in Figure 3.21 indicate that the identified dynamic-stochastic model (Equation 3.53) is adequate. Figure 3.22 shows the fresh data set and the model prediction.

Figure 3.18: Motor Load (ML) response to Transfer Screw Speed (TS)
Chapter 3. Industrial TMP Process Identification

Figure 3.19: ML/TS impulse response estimate

Figure 3.20: ML/TS residual correlations
Chapter 3. Industrial TMP Process Identification

Figure 3.21: ML/TS residual correlations (d=1)

Figure 3.22: ML/TS input, output and model prediction
Dilution Water (DW) to Motor Load (ML)

Because the primary refiner under this trial is a twin-disc refiner, two sets of PRBS signals were added to perturb the valve positions on both the drive and tail ends in order to vary the total dilution flow (Figure 3.23). To find the relationship between the motor load (ML) and total (total = drive end + tail end) dilution water (DW), ML response to the changes in the total DW is plotted out in Figure 3.24.

Figure 3.25 shows the unit impulse weights estimated using the motor load/dilution flowrate data from Figure 3.24. An increase in DW will decrease ML. This has been indicated by the first few negative impulse weights. The first and second impulse weights in Figure 3.25 are insignificant, indicating the process may have two sampling periods of pure time delay \((k = 2)\). The remaining impulse weights in Figure 3.25 suggest one of several alternatives for the orders of the transfer function polynomials.

An iterative model-building procedure consisting of structure identification, fitting, and model checking stages identified the following dynamic-stochastic model:

\[
y_t = \frac{-0.0089}{1 - 0.245z^{-1}} u_{t-3} + \frac{1 + 0.4878z^{-1}}{1 - 0.545z^{-1}} e_t
\]

Table 3.10 shows the parameter estimates, estimated standard deviations and Akaike's FPE. The residual correlations in Figure 3.20 indicate that the above identified model is adequate.

<table>
<thead>
<tr>
<th>Model (Eqn.)</th>
<th>d</th>
<th>(A(z^{-1})) (std)</th>
<th>(B(z^{-1})) (std)</th>
<th>(C(z^{-1})) (std)</th>
<th>(D(z^{-1})) (std)</th>
<th>k</th>
<th>FPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.54</td>
<td>0</td>
<td>-0.245</td>
<td>-0.0089</td>
<td>0.4878</td>
<td>-0.545</td>
<td>2</td>
<td>0.1155</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.201</td>
<td>0.002</td>
<td>0.0614</td>
<td>0.0607</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.55</td>
<td>1</td>
<td>-0.2653</td>
<td>-0.0087</td>
<td>0.3765</td>
<td>0.065</td>
<td>2</td>
<td>0.1468</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2178</td>
<td>0.0020</td>
<td>0.0535</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Forcing an integrator into the disturbance model gave:

\[
y_t = \frac{-0.0087}{1 - 0.2653z^{-1}} u_{t-3} + \frac{1 + 0.3765z^{-1}}{\Delta} e_t
\]  

(3.55)

Table 3.10 shows the parameter estimates, estimated standard deviations and Akaike's FPE. The residual correlations in Figure 3.20 indicate that the above identified model is also adequate.

![Graph showing changes in Valve Positions (VP), Dilution Water (DW) and Motor Load (ML)]

Figure 3.23: Changes in Valve Positions (VP), Dilution Water (DW) and Motor Load (ML)
Figure 3.24: Motor Load (ML) response to total Dilution Water (DW)

Figure 3.25: ML/DW impulse response estimates
Chapter 3. Industrial TMP Process Identification

Figure 3.26: ML/DW residual auto- and cross-correlations

Figure 3.27: ML/DW residual auto- and cross-correlations (d=1)
Chapter 3. Industrial TMP Process Identification

Table 3.11: Identification Results

<table>
<thead>
<tr>
<th>Model (eqn.)</th>
<th>Input</th>
<th>Output</th>
<th>Gain</th>
<th>Time Constant (sec.)</th>
<th>Time Delay (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.50</td>
<td>CP</td>
<td>ML</td>
<td>0.0461(MW/psi)</td>
<td>7.052</td>
<td>1</td>
</tr>
<tr>
<td>3.52</td>
<td>TS</td>
<td>ML</td>
<td>0.3396(MW/rpm)</td>
<td>4.995</td>
<td>8</td>
</tr>
<tr>
<td>3.54</td>
<td>DW</td>
<td>ML</td>
<td>-0.0118(MW/L/min)</td>
<td>0.711</td>
<td>2</td>
</tr>
</tbody>
</table>

3.5.3 Result Discussions

As mentioned earlier, the simplest way to find rough estimates of a multivariable relationship between inputs and outputs is to identify each input-output relationship separately. Identification results showed that wood chip refining is a dynamic-stochastic process with lower order dynamics plus time delay. To have some physical knowledge of the process, the identified discrete dynamics of the process were transformed into the continuous ones. Associated process gains, time delays and time constants are given in Table 3.11. A higher static gain between the transfer screw speed and the motor load indicates that the motor load is strongly affected by the chip feed. By comparison, the dilution water seems to have less impact on the process. Shorter time constant between the dilution flow and the motor load was identified, while longer time constants between the closing pressure, screw speed and the motor load were observed. A longer time delay between the screw speed and the motor load was identified. This was attributed to transferring wood chips to the inlet of the primary refiner. Because of the lack of an on-line consistency sensor during the mill trials, no identification was performed for the relationship between the consistency and other variables.
Chapter 3. *Industrial TMP Process Identification*

3.6 Summary and Conclusions

For control purposes, this chapter has tried to find dynamic relationships between input and output variables of a TMP plant. To avoid the complexity of developing differential equations in mechanistic modelling, process identification (an experimental method) was used to find rough estimates of the dynamics and disturbances of the process. Identification results indicated that wood chip refining is a dynamic-stochastic process with excessive stochastic disturbances. Due to product quality concerned, the input perturbation signal was limited to a certain range therefore signal to noise ratio was limited. For this particular case, the model was limited to lower order such as second order in order to obtain rough estimates of process dynamics with representational adequacy. It was found that first-order dynamics plus time delay gave more adequate approximation in terms of minimizing Akaike’s FPE and residual cross-correlations. As mentioned earlier, the simplest way to find estimates of a multivariable relationship between inputs and outputs is to identify the relationship for each input-output pair separately. Figure 3.28 shows identified multivariable dynamics of the process. As can be seen from mechanistic modelling, the process has nonlinear and time-varying characteristics. Therefore, the process may be described by its dynamics with unit gain plus the nonlinear relationship obtained from mechanistic modelling stage.

Figure 3.28: Dynamic and static relationships between inputs and outputs

\[
\text{disturbances} \\
\text{CP} \quad \text{TS} \quad \text{DW} \quad \text{1st order + delay} \quad \text{N. L. Mech. model} \\
\text{ML} \quad \text{Cons}
\]

CP: hydraulic closing pressure  
TS: chip transfer screw speed  
DW: dilution water flowrate  
ML: motor load  
Cons: consistency

For control purposes, this chapter has tried to find dynamic relationships between input and output variables of a TMP plant. To avoid the complexity of developing differential equations in mechanistic modelling, process identification (an experimental method) was used to find rough estimates of the dynamics and disturbances of the process. Identification results indicated that wood chip refining is a dynamic-stochastic process with excessive stochastic disturbances. Due to product quality concerned, the input perturbation signal was limited to a certain range therefore signal to noise ratio was limited. For this particular case, the model was limited to lower order such as second order in order to obtain rough estimates of process dynamics with representational adequacy. It was found that first-order dynamics plus time delay gave more adequate approximation in terms of minimizing Akaike’s FPE and residual cross-correlations. As mentioned earlier, the simplest way to find estimates of a multivariable relationship between inputs and outputs is to identify the relationship for each input-output pair separately. Figure 3.28 shows identified multivariable dynamics of the process. As can be seen from mechanistic modelling, the process has nonlinear and time-varying characteristics. Therefore, the process may be described by its dynamics with unit gain plus the nonlinear relationship obtained from mechanistic modelling stage.

Figure 3.28: Dynamic and static relationships between inputs and outputs
Chapter 4

Constrained Multivariable Model-Based Predictive Control

4.1 Introduction

Model-Based Predictive Control (MBPC) refers to a class of algorithm that computes a sequence of manipulated variable adjustments in order to optimize the future behavior of a process. Originally developed to meet the specialized control needs of power plants and petroleum refineries, MBPC technology is probably the most important approach to the advanced control of complex interacting industrial processes[78] and is gaining widespread acceptance in various industries. MBPC’s effectiveness in practice was discovered in the early 80’s by the development and application of model predictive heuristic approaches: IDCOM (Identification/Command) by Richalet et al.[79] and DMC (Dynamic Matrix Control) by Cutler and Ramaker [80]. In both algorithms an explicit dynamic model of a plant is used to predict the effect of future actions of the manipulated variables on the output (thus the name "Model Predictive Control"). The future control actions are determined by minimizing the predicted error subject to operating constraints. The calculation of future control actions is repeated at each sampling time based on updated measurements from the plant. A recent paper written by Froisy [81] provides an interesting industrial perspective of current MBPC technology and summarizes likely future developments. Clarke and Zhang [82] compared the different MBPC approaches and found them all suited to plants that exhibit variable dead-time and non-minimum phase characteristics. A comprehensive survey of the history and theory of
model-based controls can be found in various papers [83, 84, 85, 86, 87].

Generalized Predictive Control (GPC) first introduced by Clarke et. al. [88] is one of a class of MBPC algorithms. Few advanced control methods have had as much influence, widespread acceptance and success in industrial applications as the GPC approach. The success of this technique is due to its capabilities of controlling a process with:

- variable time delay and model order;
- over-parameterization (plant/model mis-match);
- unstable zeros (non-minimum phase);
- unstable poles;
- load-disturbances.

In addition, GPC can successfully overcome difficulties encountered by other control methods. Minimum-variance self-tuning has been used for a long time, but it is highly sensitive to the assumption made about the dead-time. Pole-placement and Linear Quadratic Gaussian (LQG) self-tuners perform badly if the order of a process is overestimated because of pole/zero cancellations in the identified model. Although self-tuning and adaptive controls have made much progress over the previous decade in terms of theoretical understanding and practical applications, neither one is suitable as a general purpose algorithm for the stable control of the majority of real processes [88]. The GPC algorithm possesses simplicity and flexibility, appearing to overcome the problems in one algorithm. As far as process constraints are concerned, GPC has the capability of incorporating and handling the constraints explicitly. However, relatively little information is available on constraint handling using other methods.

The refiner mechanical pulping process under this study is a rather complex process. Good control will improve the production, efficiency and pulp quality, while poor control
can lead to operating difficulties and poor physical and papermaking properties. Developing a suitable control strategy for the process is a very challenging task, because of the multivariable nature, nonlinearity, process interactions, time varying, time-delay, variable constraints and stochastic disturbances of the process. A controller for the process must be able to handle the problems together with the robustness to modelling errors. Constrained multivariable adaptive-predictive control approach presented in this chapter will be used for the control of the process. The controller is developed based on the GPC approach, where future control actions are determined by minimizing the predicted error subject to input-output constraints. A process model is used to predict future process output under the influence of future control actions. The model parameters are estimated at each sampling time based on update measurements from the process. Quadratic programming and mixed-weights least-squares algorithms are introduced to solve the constrained control problem. The proposed controller will possess a number of important features as follows:

- The controller is able to handle process/model mismatch (inaccurate time-delay and model order) and disturbances.

- The controller adopts an integrator as a natural consequence of its assumption about the basic process model to obtain offset-free close-loop behavior of the control.

- An embedded adaptive scheme can help to handle process nonlinear and time-varying characteristics under the help of a proper parameter estimation method.

- An explicit model is used in the controller development so that fewer parameters are needed to be estimated on-line than equivalent implicit schemes and thus it could be very beneficial to the parameter estimation of an MIMO model.
The proposed predictive control strategy can help to handle variable constraints explicitly and naturally by taking into account the constraints into the controller development.

This chapter is structured as follows. Section 4.2 gives a brief introduction to unconstrained Generalized Predictive Control (UCGPC) for both single-input single-output (SISO) and multi-input multi-output (MIMO) cases. Section 4.3 describes recursive parameter estimation methods that may incorporate with an adaptive control scheme to compensate for the time-varying and nonlinear features of the process. Section 4.4 presents a predictive control problem subject to input and output constraints. In Section 4.5, an analytical solution of the constrained GPC is derived by solving analytical quadratic programming (QP) problem. Section 4.6 presents the derivation of the constrained GPC for the general cases by solving a mixed-weights least-squares (MWLS) problem. The associated weights in MWLS can be modified in a manner that encompasses both the requirement for the future control moves to lie inside the feasible region and minimization of desired control objective function. Simulations of the proposed control strategy are presented in Section 4.6. Summary and conclusions are given in Section 4.7.

4.2 Unconstrained Generalized Predictive Control

4.2.1 Basic SISO Generalized Predictive Control

Model and Output Prediction

A model is the center for any kind of model-based control designs. The model used in GPC design is the Controlled Autoregressive and Integrated Moving Average (CARIMA)
model as:

\[ A(z^{-1})y(t) = B(z^{-1})u(t - k) + \frac{C(z^{-1})}{\Delta} \xi(t) \]  

(4.56)

where \( y(t) \) is the process output, \( u(t) \) is the input (\( y \) and \( u \) are deviation variables), and \( \xi(t) \) is uncorrelated, normally distributed zero-mean random variable (or white noise) with the variance \( \sigma^2_\xi \). \( k \) is the time delay. The difference operator \( \Delta = 1 - z^{-1} \) in the denominator of the noise term is widely assumed, as it forces an integrator into the controller in order to eliminate offset between the measured output and its setpoint. \( A(z^{-1}), B(z^{-1}) \) and \( C(z^{-1}) \) are the polynomials in the backward-shift operator \( z^{-1} \) with the orders of \( na, nb \) and \( nc \) respectively:

\[
A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{na} z^{-na} \\
B(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{nb} z^{-nb} \\
C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_{nc} z^{-nc} 
\]  

(4.57)

The above CARIMA model has a more general application. It has also been used by Tuff and Clarke to derive the generalized minimum variance control (GMV) and pole-placement self-tuners with inherent integral action [89].

To derive a \( j \)-step ahead predictor \( \hat{y}(t+j|t) \) from Equation 4.56, consider the following Diophantine identity:

\[
C(z^{-1}) = E_j(z^{-1})A(z^{-1}) \Delta + z^{-j}F_j(z^{-1}) 
\]  

(4.58)

\( E_j(z^{-1}) \) and \( F_j(z^{-1}) \) (order \( j - 1 \) and \( \max(na, nc - j) \), respectively) are determined by the given polynomials \( A(z^{-1}), C(z^{-1}) \) and the prediction interval \( j \). Note that the disturbance in Equation 4.56 is modelled as a nonstationary process by \( \Delta(z^{-1}) \). If the model is obtained using identification techniques, \( C(z^{-1}) \) is the polynomial affected by a bigger error. Therefore, in GPC, the \( C(z^{-1}) \) polynomial is considered as a design polynomial chosen by the user to obtain good behavior of the controller [88].
Multiplying Equation 4.56 by $E_j(z^{-1}) \triangle z^j$ and combining with Equation 4.58 gives:

$$y(t + j) = G_j(z^{-1}) \triangle u^f(t + j - 1) + F_j(z^{-1})y^f(t) + E_j(z^{-1})\xi(t + j) \quad (4.59)$$

where $G_j(z^{-1}) = E_j(z^{-1})B(z^{-1})$ in the order of $j + nb - 1$. Note, $k - 1$ leading zeros have been added to $B(z^{-1})$ in Equation 4.56 to simplify the notation. Hence the polynomial $B(z^{-1})$ is in the order of $nb = nb + k - 1$. The superscript $'$ in the above equation denotes filtering by $1/C(z^{-1})$. As the polynomial $E_j(z^{-1})$ is of the degree $j - 1 (j \geq 1)$, the noise terms $E_j(z^{-1})\xi(t + j)$ in the above equation are all future unknown values and thereby the output prediction in Equation 4.59 can be modified as follows:

$$\hat{y}(t + j|t) = G_j(z^{-1}) \triangle u^f(t + j - 1) + F_j(z^{-1})y^f(t) \quad (4.60)$$

To separate past known filtered control actions from current control actions yet to be determined, consider the following identity:

$$G_j(z^{-1}) = G_j'(z^{-1})C(z^{-1}) + z^{-j}\Gamma_j(z^{-1}) \quad (4.61)$$

where $G_j'(z^{-1})$ and $\Gamma_j(z^{-1})$ (order $j - 1$ and max$(nb - 1, nc - 1)$, respectively) are defined by $G_j(z^{-1}), C(z^{-1})$ and $j$. Substituting the above equation into Equation 4.60 gives:

$$\hat{y}(t + j|t) = G_j'(z^{-1}) \triangle u(t + j - 1) + \Gamma_j(z^{-1}) \triangle u^f(t - 1) + F_j(z^{-1})y^f(t) \quad (4.62)$$

where

$$f(t + j) = \Gamma_j(z^{-1}) \triangle u^f(t - 1) + F_j(z^{-1})y^f(t) \quad (4.63)$$

As can be seen from the above equation, the output prediction $\hat{y}(t + j|t)$ consists of two types of process responses:

- $G_j'(z^{-1}) \triangle u(t + j - 1)$ is called "forced response" of the process. It depends on the future control actions yet to be determined.
• \( f(t + j) \) is called "free response" of the process. It depends on the past filtered known controls together with filtered measured outputs under the assumption that no future control actions are performed.

Predictive Control Law

The control law of GPC for a SISO case is derived from a quadratic optimal cost function or the function based on a designed trade-off between the control performance and soft constraints on the control actions as follows:

\[
J_{GPC} = E \left\{ \sum_{j=N_1}^{N_2} [y(t+j) - w(t+j)]^2 + \sum_{j=1}^{NU} \rho(j) \Delta u^2(t+j-1) \right\} \quad (4.64)
\]

where

- \( y(t+j) \): the future output
- \( w(t+j) \): the future setpoint
- \( u(t+j-1) \): the future control action
- \( N_1 \): the minimum predictive horizon, \((N_1 \geq 1)\)
- \( N_2 \): the maximum prediction horizon, \((N_2 \geq N_1)\)
- \( NU \): the control horizon, \((N_2 \geq NU \geq 1)\)
- \( \rho(j) \): the control-weighting factor

The objective of the predictive control is to drive future plant output \( y(t+j) \) close to the future setpoint \( w(t+j) \) in some sense by selecting a control movement so that the desired performance index of Equation 4.64 is minimized. The future plant output \( y(t+j) \) is predicted based on the plant model with knowing the current and past control actions and past plant output.

To simplify the following derivation, the control weighting factor and the minimum predictive horizon are set to \( \rho(j) = \rho \) and \( N_1 = 1 \), respectively. The objective function
Chapter 4. Constrained Multivariable Model-Based Predictive Control

(Equation 4.64) and the output prediction (Equation 4.62) can then be stated in more compact forms as:

\[ J_{GPC} = (y - w)^T(y - w) + \rho \tilde{u}^T\tilde{u} \tag{4.65} \]

and

\[ \hat{y} = G'\tilde{u} + f \tag{4.66} \]

where

\[ \hat{y} = [\hat{y}(t + 1), \hat{y}(t + 2), \ldots, \hat{y}(t + N_2)]^T \]

\[ \tilde{u} = [\Delta u(t), \Delta u(t + 1), \ldots, \Delta u(t + NU - 1)]^T \tag{4.67} \]

\[ f = [f(t + 1), f(t + 2), \ldots, f(t + N_2)]^T \]

\[ w = [w(t + 1), w(t + 2), \ldots, w(t + N_2)]^T \]

and the matrix \( G' \) is lower-triangular of dimension \( N_2 \times NU \):

\[ G' = \begin{bmatrix}
  g_0' & 0 & \cdots & 0 \\
  g_1' & g_0' & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  g_{NU-1}' & \cdots & g_1' & g_0' \\
  \vdots & \vdots & \ddots & \vdots \\
  g_{N_2-1}' & g_{N_2-2}' & \cdots & g_{N_2-NU}'
\end{bmatrix} \]

Substituting Equation 4.66 into 4.65 and minimizing with respect to \( \tilde{u} \) yields:

\[ \tilde{u} = (G'^T G' + \rho I)^{-1} G'^T(w - f) \tag{4.68} \]

As GPC is a receding-horizon control strategy, only the first control increment \( \Delta u(t) \) in the control vector \( \tilde{u} \) is implemented into the plant under the assumption that beyond the control horizon \( NU \), further increments in control actions are zero, i.e., \( \Delta u(t + NU) = \Delta u(t + NU + 1) = \cdots = \Delta u(t + N_2) = 0 \) (or \( u(t + NU - 1) = u(t + NU) = u(t + NU + 1) = \cdots \)).
\( \cdots, = u(t+N_2) \). At the time \( t+1 \), the computation is repeated with the horizon moved by one time interval. The GPC control algorithm described above represents an extremely flexible approach for the control of different processes due to its various design and tuning parameters incorporated into the algorithm. The role of setting these design and tuning parameters is to obtain a satisfactory closed-loop performance for different processes.

**Predictive horizon** \((N_1, N_2)\): predicting the output response over the predictive horizon enables compensation for any variable or unknown time delays. The minimum predictive horizon \(N_1\) can often be taken as 1 if the plant dead-time is unknown. A small value of the maximum predictive horizon \(N_2\) will result in more active control effort and the performance is quite similar to the performance of a minimum-variance control. While with a larger value of \(N_2 \to \infty\), the control performance will be very sluggish.

**Control horizon** \((NU)\): a decrease in the control horizon \(NU\) will significantly reduce the computational effort and simplify control derivation, while an increase in \(NU\) will result in excessive movements in the control action and plant output.

**Control weighting** \((\rho)\): it is designed to reduce the control movements. Zero control weighting (i.e. \(\rho = 0\)) gives a dead-beat response, while with a very large value of \(\rho\), the control action is dampened significantly.

To reduce computation complexity, it is always advisable to take prior knowledge of plant delay and order into account so that \(N_1\) and \(N_2\) can be selected. It is also possible to reduce the computational burden by imposing a constant control input after some control horizon \(NU\). It is found that a large class of plant models can be stabilized by GPC with \(N_1 = 1, N_2 = 10\) [88] and that it will give an acceptable control by setting \(NU = 1\) for simple processes. The details of guidelines for the principal design parameters and stabilization theorems are available in [90, 91, 92].
4.2.2 Basic MIMO Generalized Predictive Control

In the following, the development of an MIMO GPC is a direct extension of the SISO GPC described in the previous section. The derivation of a multivariable GPC can also be found in the papers by Shah et al.[93] and Mohtadi et al.[94]. The following is a brief description of the derivation of the MIMO GPC.

Model and Output Prediction

Consider an MIMO CARIMA model for a m-input n-output process as follows:

\[ A(z^{-1}) \triangle y(t) = B(z^{-1}) \Delta u(t - 1) + C(z^{-1})e(t) \]  

(4.69)

where

- \( u(t), y(t) \) and \( e(t) \) are the vectors of the process inputs, outputs and disturbances of the dimension \( m, n \) and \( n \) respectively.
- \( A \) and \( C \) are diagonal polynomial matrices of the dimension \( n \times n \), while \( B \) is a polynomial matrix of the dimension \( n \times m \).

The time delay of the process has been included in \( B \) matrix. If any of the input-output pairs of the process has a non-zero dead-time, the leading elements of the corresponding polynomial in \( B \) are zeros.

To derive a \( j \)-step ahead predictor \( \hat{y}(t + j|t) \), the following matrix Diophantine equation is considered:

\[ C(z^{-1}) = E_j(z^{-1})A(z^{-1}) \triangle + z^{-j}F_j(z^{-1}) \]  

(4.70)

where \( E_j(z^{-1}) \) and \( F_j(z^{-1}) \) are diagonal polynomial matrices, defined by \( A(z^{-1}) \) and \( C(z^{-1}) \). \( E_j(z^{-1}) \) is has a degree of \( j - 1 \). Multiplying Equation 4.69 by \( E_j(z^{-1})z^j \) and
combining with Equation 4.70 yields:

\[ y(t + j) = G_j(z^{-1}) \Delta u^f(t + j - 1) + F_j(z^{-1})y^f(t) + E_j(z^{-1})e(t + j) \]  

(4.71)

where \( G_j = E_jB \). The superscript \( f \) denotes filtering by \( C(z^{-1})^{-1} \). As \( E_j(z^{-1}) \) is a degree of \( j - 1 (j \geq 1) \), the noise terms \( E_j(z^{-1})e(t + j) \) are all future unknown values. Therefore, the above output predictor can be modified as:

\[ \hat{y}(t + j|t) = G_j(z^{-1}) \Delta u^f(t + j - 1) + F_j(z^{-1})y^f(t) \]  

(4.72)

To separate the past known filtered control actions from the current control actions yet to be determined, the following polynomial matrix identity is considered:

\[ G_j(z^{-1}) = G'_j(z^{-1})C(z^{-1}) + z^{-j}L_j(z^{-1}) \]  

(4.73)

Substituting Equation 4.73 into 4.72 gives:

\[ \hat{y}(t + j|t) = G'_j(z^{-1}) \Delta u(t + j - 1) + L_j(z^{-1}) \Delta u^f(t - 1) + F_j(z^{-1})y^f(t) \]  

(4.74)

where

\[ f(t + j) = L_j(z^{-1}) \Delta u^f(t - 1) + F_j(z^{-1})y^f(t) \]  

(4.75)

**MIMO Predictive Control Law**

The optimal movements of control actions are selected such that the following desired performance index is minimized:

\[ \min_{\mathbf{u}} J = \sum_{j=N_1}^{N_2}[y(t + j) - w(t + j)]^T R[y(t + j) - w(t + j)] + \sum_{j=1}^{N_U} \Delta u^T(t + j - 1)Q \Delta u(t + j - 1) \]  

(4.76)

where
• $R$ is the output weighting matrix with positive elements on its diagonal, i.e., $R = \text{diag}\{r_1, r_2, \ldots, r_n\}$.

• $Q$ is the control weighting matrix with positive elements on its diagonal, i.e., $Q = \text{diag}\{q_1, q_2, \ldots, q_m\}$.

• $w(t+j)$ is the output setpoint vector of dimension $n$.

Substituting Equation 4.74 into 4.76 gives:

$$
\min_{\mathbf{u}} J = [\tilde{G}\mathbf{u} + \mathbf{f} - \mathbf{w}]^T R [\tilde{G}\mathbf{u} + \mathbf{f} - \mathbf{w}] + \mathbf{u}^T Q \mathbf{u} \quad (4.77)
$$

where

$$\begin{align*}
\mathbf{u} &= [\Delta \mathbf{u}(t) \quad \Delta \mathbf{u}(t+1) \cdots \Delta \mathbf{u}(t+NU-1)]^T \\
\mathbf{f} &= [f_1(t+1) \cdots f_n(t+1) \cdots f_1(t+N_2) \cdots f_n(t+N_2)]^T \\
\mathbf{w} &= [w_1(t+1) \cdots w_n(t+1) \cdots w_1(t+N_2) \cdots w_n(t+N_2)]^T
\end{align*}
$$

and $\tilde{G}$ is the matrix of dimension $n(N_2) \times m(NU)$ in the case of $N_1 = 1$:

$$
\tilde{G} = \begin{bmatrix}
G'_0 & 0 & \cdots & 0 \\
G'_1 & G'_0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
G'_{NU-1} & \cdots & G'_1 & G'_0 \\
\vdots & \vdots & \ddots & \vdots \\
G'_{N2-1} & G'_{N2-2} & \cdots & G'_{N2-NU}
\end{bmatrix} \quad (4.79)
$$

Minimizing Equation 4.77 with respect to $\mathbf{u}$ yields:

$$
\Delta \mathbf{u}(t) = [I_m \ 0 \cdots ] [\tilde{G}^T R \tilde{G} + Q]^{-1} \tilde{G} (\mathbf{w} - \mathbf{f}) \quad (4.80)
$$
4.3 Recursive Parameter Estimation in Adaptive Control

As mentioned previously, the process under this study possesses time-varying and non-linear characteristics. The best solution to compensate for the time-varying nature is to identify the changes and make an adjustment in the process model. This leads us to considering an adaptive version of a model predictive control incorporating a proper parameter estimation method. By selecting an adaptive scheme, the nonlinear problems can be treated as a linear one with a time-varying gain. Therefore, a more complex nonlinear control algorithm can be avoided.

Parameter estimation incorporated with an adaptive control includes various recursive identification algorithms [95]. The most wildly used scheme of these algorithms is the Recursive Least Squares (RLS). The algorithm is:

\[
\begin{align*}
\hat{e}_t & = y_t - \phi_t^T \hat{\theta}_{t-1} \\
\hat{\theta}_t & = \hat{\theta}_{t-1} + P_t \phi_t \hat{e}_t \\
P_t & = \lambda^{-1} \left[ I - \frac{P_{t-1} \phi_t \phi_t^T}{\lambda + \phi_t^T P_{t-1} \phi_t} \right] P_{t-1}
\end{align*}
\]  

(4.81)

where \( \phi_t \) is the vector of past inputs and outputs, \( \hat{\theta}_t \) is the vector of the model parameters estimated, \( P_t \) is the square matrix proportional to the covariance of the parameters, \( \hat{e}_t \) is the prediction error and \( \lambda \) is the forgetting factor.

The basic RLS is known to have optimal properties when the parameters are time invariant, but it is unsuitable for tracking time-varying parameters since the algorithm gain converges to zero. Modified versions of the algorithm have been developed under the effort of many researchers. An improved algorithm is Exponential Forgetting and Resetting Algorithm (EFRA) algorithm proposed by Salgado et al. [96]. The algorithm is expressed as:

\[
\hat{e}_t = y_t - \phi_t^T \hat{\theta}_{t-1}
\]
\[ \dot{\theta}_t = \hat{\theta}_{t-1} + \frac{\alpha P_{t-1}^{i} \hat{\phi}_t}{1 + \hat{\phi}_t^T P_{t-1} \hat{\phi}_t} \hat{e}_t \] (4.82)
\[ P_t = \frac{1}{\lambda} P_{t-1} - \frac{\alpha P_{t-1} \hat{\phi}_t^T P_{t-1}}{1 + \hat{\phi}_t^T P_{t-1} \hat{\phi}_t} + \beta I - \delta P_{t-1}^2 \]

where the estimator’s parameters \( \alpha, \beta, \delta \) and \( \lambda \) are constants. By making particular choices for \( \alpha, \beta, \delta \) and \( \lambda \), the algorithm is reduced to many of the previous RLS based algorithms. Typical values of \( \alpha, \beta, \delta \) and \( \lambda \) suggested in [96] are \( \alpha \in [0.1, 0.5], \beta \in [0, 0.01], \delta \in [0, 0.01] \) and \( \lambda \in [0.9, 0.99] \). With these parameter ranges, the covariance matrix \( P_t \) is bounded to \( P_{\text{min}} \leq P_t \leq P_{\text{max}} \). This algorithm incorporates both exponential forgetting old data to ensure continual parameter adaptation and exponential covariance resetting during periods of low excitation. Complete discussions on various parameter estimation algorithms for the linear regression type model can be found in Ljung [77] and Åstrom and Wittenmark [97].

### 4.4 Control under Constraints

Constraints that are strictly enforced are referred to as hard constraints. Hard constraints are often taken into account when undesirable and unrealistic control inputs are computed because of process changes. These changes may happen during setpoint change or more generally when the process dynamic varies and leads to the control inputs which cannot be actually applied due to the reasons of safety, equipment physical limitations or product quality requirement. In addition, the economic operating condition of a typical process unit often lies at the intersection of constraints [98]. A successful industrial controller should, therefore, be able to maintain the process as close as possible to the constraints but without violating them.

The unconstrained control criterion described earlier is to calculate a sequence of future control actions by minimizing a desired performance index. In the case when the
controller which gives the minimum variance requires too large a control signal, it will result in the process operating out of its pre-specified operating limits. Although the control actions can be reduced by choosing relatively high control weights in the GPC algorithm, it may not be easy to find an accurate value to guarantee pre-specified variable constraints. This motivated the current study on the development of a constrained control strategy that would be suitable for the control of a TMP plant.

The typical constraints for a SISO case may include amplitude/rate limits on the control signal and amplitude limits on the process output, i.e.

\[ u_{\text{min}} \leq u(t + i - 1) \leq u_{\text{max}} \]
\[ \Delta u_{\text{min}} \leq \Delta u(t + i - 1) \leq \Delta u_{\text{max}} \]
\[ y_{\text{min}} \leq y(t + j) \leq y_{\text{max}} \]
\[ i = 1, \ldots, NU \]
\[ j = N_1, \ldots, N_2 \]

where

\[ (u_{\text{min}}, u_{\text{max}}) \] : the lower and upper bounds on the control amplitude.
\[ (\Delta u_{\text{min}}, \Delta u_{\text{max}}) \] : the lower and upper bounds on the control increment.
\[ (y_{\text{min}}, y_{\text{max}}) \] : the lower and upper bounds on the process output amplitude.

In GPC-based control development, the process is assumed to be linear. Hence any constraints on the future inputs and outputs (Equation 4.83) can be mapped to a linear combination of the future control increments according to:

\[ u_{\text{min}} \leq u(t + i - 1) \leq u_{\text{max}} \]
\[ \Downarrow \]
\[ u_{\text{min}} - u(t + i - 2) \leq \Delta u(t + i - 1) \leq u_{\text{max}} - u(t + i - 2) \]
and

\[
\begin{bmatrix}
\begin{array}{c}
\mathbf{y}_{\min} \\
\vdots \\
\mathbf{y}_{\min}
\end{array}
\end{bmatrix}
\leq
\begin{bmatrix}
\begin{array}{c}
y(t + N_1) \\
y(t + N_1 + 1) \\
\vdots \\
y(t + N_2)
\end{array}
\end{bmatrix}
\leq
\begin{bmatrix}
\begin{array}{c}
\mathbf{y}_{\max} \\
\vdots \\
\mathbf{y}_{\max}
\end{array}
\end{bmatrix} = \mathbf{y}_{\max}
\]

However, it must be kept in mind that the mapping of process constraints depends greatly on the accuracy of the process model. Only an accurate model will result in the required effect as mapping of output constraints to the constraints on control increments \( \ddot{\mathbf{u}} \).

Substituting Equations 4.84 and 4.85 into Equation 4.83 gives a more compact form as follows (see Appendix C for the details):

\[
\begin{bmatrix}
A_1 \\
-A_1 \\
I_{NU \times NU} \\
-I_{NU \times NU} \\
\mathbf{G}' \\
-\mathbf{G}'
\end{bmatrix} \ddot{\mathbf{u}} \leq
\begin{bmatrix}
\mathbf{u}_{\max} - A_2 u(t - 1) \\
-\mathbf{u}_{\min} + A_2 u(t - 1) \\
\triangle \mathbf{u}_{\max} \\
- \triangle \mathbf{u}_{\min} \\
\mathbf{y}_{\min} - \mathbf{f} \\
- \mathbf{y}_{\max} + \mathbf{f}
\end{bmatrix}
\]

where

\[
\ddot{\mathbf{u}} = [u(t) \ u(t + 1) \ldots u(t + NU - 1)]^T
\]
and

\[
A_1 = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
1 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \cdots & 0
\end{bmatrix}
\]

(4.88)

\[A_2 = [1 \ 1 \cdots 1]^T\]

The elements in the vectors \(u_{\text{max}}, u_{\text{min}}, \Delta u_{\text{max}}, \Delta u_{\text{min}}, y_{\text{max}}\) and \(y_{\text{min}}\) are associated values of the input and output constraints.

If none of the constraints in Equation 4.86 is violated, the control actions derived by minimizing a desired performance index alone are the optimal solution. If however one or more constraints in Equation 4.86 are violated, the solution will no longer be optimal. The optimal solution should be obtained by minimizing the performance index with satisfaction of the input-output constraints, i.e.

\[
\min_{\hat{u}} J = [G'\hat{u} + f - w]^T R [G'\hat{u} + f - w] + \hat{u}^T Q \hat{u}
\]

subject to:

\[
\begin{bmatrix}
A_1 \\
-A_1 \\
I_{N_u \times N_u} \\
-I_{N_u \times N_u} \\
G' \\
-G'
\end{bmatrix}
\begin{bmatrix}
\hat{u}
\end{bmatrix}
\leq
\begin{bmatrix}
\Delta u_{\text{max}} \\
-\Delta u_{\text{min}} \\
y_{\text{min}} - f \\
y_{\text{max}} + f
\end{bmatrix}
\]

(4.89)

The similar mathematical formula can be derived for MIMO cases. The conceptual structure of the above constrained model-based predictive control strategy is also illustrated in Figure 4.29.
There is no analytical solution available for the general case of constrained predictive control [99]. Some analytical optimal solutions exist only for some special cases [100, 99]. The solution for a general case may be obtained with the help of some available commercial optimization software packages, but it turns out to be computationally demanding and may not be suitable for on-line implementations. The following presents the derivation of an analytical solution for the constrained GPC for a simple case. An optimal solution of a constrained GPC for the general case is derived by solving a mixed-weights least-squares (MWLS) problem.

4.5 Constrained Control via Quadratic Programming

An optimal solution of a constrained control problem was first developed by using linear programming, but the solution of the constrained problem was obtained using quadratic programming [101]. There are several algorithms that exist for solving QP problems. They can be roughly divided into two groups: those based on the computation of Kuhn
Tucker multipliers [102] and Lagrange multipliers [103], and those reducing a QP problem to non-negative least squares [104] or to a linear complementary problem [105, 103]. Constrained SISO GPC for the case of $NU \leq 2$ was treated by Tsang and Clarke [100] using the Lawson and Hanson method [104], where the control signal for the rate or amplitude constraint (but not both) was calculated. Tsang and Clarke concluded in [100] that the Lawson and Hanson method is rather complex and computationally demanding, so it is not adequate for fast processes application. Camacho [106] transformed the GPC QP problem to a linear complementarity problem and solved it using the Lemke’s algorithm. In Camacho’s work, a simple algorithm to decrease the number of constraints and a modification to a standard algorithm for solving the linear complementary problem was used in order to reduce the amount of computation required. Recently, an analytical solution to the constrained GPC was derived by Mutha [99] through the computation of Kuhn Tucker multipliers for $NU \leq 2$ in a SISO GPC case and for $NU = 1$ in a MIMO GPC case. However, Mutha’s work did not take output constraints into consideration in the controller development. In the following, an analytical solution to the constrained GPC subject to both input and output constraints is derived by modifying Mutha’s method.

4.5.1 Quadratic Programming via Kuhn Tucker Multipliers Computation

Kuhn Tucker multipliers method can be used to solve the minimization of a nonlinear quadratic cost function with linear constraints as shown below [102]:

$$\min f(x)$$

subject to:

$$h_i(x) \leq 0 \text{ for } i = 1 \text{ to } p$$

$$g_j(x) = 0 \text{ for } j = 1 \text{ to } q$$
where \( f(x) \) is assumed to be continuous and continuously differentiable up to 1st (for the 1st order conditions) or 2nd order (for the 2nd order conditions) in the feasible region \( D \) which is defined by \( p + q \) constraints. \( x \in R^m \) is said to be the "regular point" of the constraints if at this point all the constraints are independent. The above optimal problem can be rewritten as follows by including the constraints as an unconstrained problem:

\[
\Phi(x, \lambda, \mu) = f(x) + \sum_{j=1}^{q} \lambda_j g_j(x) + \sum_{i=1}^{p} \mu_i h_i(x)
\]  

(4.92)

where \( \mu_i \) are Kuhn-Tucker multipliers and \( \lambda_j \) are Lagrangian multipliers. For \( \Phi \) to be minimum, the following 1st order and 2nd order conditions must be satisfied.

The 1st order conditions are:

\[
\frac{\partial \Phi}{\partial x} = 0
\]

\[
\frac{\partial \Phi}{\partial \lambda} = 0
\]

\[\mu^T h(x) = 0\]  

(4.93)

The 2nd order conditions are:

\[\mu_i \geq 0 \text{ for } i = 1 \text{ to } p\]  

(4.94)

\[\frac{\partial^2 \Phi}{\partial^2 x} \text{ is non-negative definite in the subspace of the constraints}\]

Note that the sign of the Kuhn-Tucker multipliers \( \mu_i \) are subject to change with the nature of the problem (see Table 4.12 for different cases).

The above conditions do not enable us to compute the values \( x \) directly since they are in general nonlinear and unsolvable analytically. Conventional solutions for the problem may be via Iterative Search Methods [102], but it turns out to be computationally demanding. Hence the use of an analytical solution to solve the above problem is highly desirable as it requires reduced computation.
Nature of the constraints | Nature of the extremes
---|---
Max \( f(x) \) | \( \mu_i \leq 0 \)
Min \( f(x) \) | \( \mu_i \geq 0 \)
subject to \( h_i(x) \leq 0 \) | \( \mu_i \leq 0 \)
subject to \( h_i(x) \geq 0 \) | \( \mu_i \geq 0 \)

Table 4.12: Kuhn-Tucker Multipliers for Different Cases

### 4.5.2 Analytical Solution for Constrained SISO GPC

Constrained SISO GPC for \( NU = 1 \) is described as below:

\[
\min_{\Delta u(t)} J = \left[ G' \Delta u(t) + \mathbf{f} - \mathbf{w} \right]^T \left[ G' \Delta u(t) + \mathbf{f} - \mathbf{w} \right] + \rho \Delta u^2(t)
\]

subject to:

\[
u_{\text{min}} \leq u(t) \leq u_{\text{max}}
\]

\[
\Delta u_{\text{min}} \leq \Delta u(t) \leq \Delta u_{\text{max}}
\]

\[
y_{\text{min}} = \begin{bmatrix} y_{\text{min}} \\ y_{\text{min}} \\ \vdots \\ y_{\text{min}} \end{bmatrix} \leq \begin{bmatrix} y(t + N_1) \\ y(t + N_1 + 1) \\ \vdots \\ y(t + N_2) \end{bmatrix} \leq \begin{bmatrix} y_{\text{max}} \\ y_{\text{max}} \\ \vdots \\ y_{\text{max}} \end{bmatrix} = y_{\text{max}}
\]

where \( y_{\text{min}} = [y_{\text{min}}, y_{\text{min}}, \ldots, y_{\text{min}}]^T \) and \( y_{\text{max}} = [y_{\text{max}}, y_{\text{max}}, \ldots, y_{\text{max}}]^T \) are the vectors including the lower and higher bounds of the process output.

In GPC-based control, as stated earlier, the process is assumed to be linear. Hence the input and output constraints in the above equation can be mapped as a linear combination of the control increment \( \Delta u(t) \) by substituting \( u(t) = \Delta u(t) + u(t - 1) \) and \( y = G' \Delta u(t) + \mathbf{f} \) into the above equation, i.e.:

\[
u_{\text{min}} - u(t - 1) \leq \Delta u(t) \leq u_{\text{max}} - u(t - 1)
\]

\[
\Delta u_{\text{min}} \leq \Delta u(t) \leq \Delta u_{\text{max}}
\]

\[
\frac{G'T(y_{\text{min}} - \mathbf{f})}{G'TG'} \leq \Delta u(t) \leq \frac{G'T(y_{\text{max}} - \mathbf{f})}{G'TG'}
\]

As the optimal problem in Equation 4.95 involves only one variable \( \Delta u(t) \), the minimization of the cost function subject to the constraints is solved by simply clipping the
unconstrained solution $\Delta u^*(t)$ to the appropriate bounds, i.e.

$$
\Delta u(t) = \begin{cases} 
  a & \text{if } \Delta u^*(t) < a \\
  \Delta u^*(t) & \text{if } a \leq \Delta u^*(t) \leq b \\
  b & \text{if } b < \Delta u^*(t)
\end{cases}
$$

(4.97)

where $a$ and $b$ are the scalars associated with the constraints, and given by:

$$
a = \max \left\{ \Delta u_{\min}, u_{\min} - u(t - 1), \frac{G'G}{G'TG'} (y_{\min} - f) \right\}
$$

$$
b = \min \left\{ \Delta u_{\max}, u_{\max} - u(t - 1), \frac{G'G}{G'TG'} (y_{\max} - f) \right\}
$$

(4.98)

The analytical solution for $NU = 2$ in a SISO GPC case can be derived similarly as the derivation of the analytical solution for a constrained MIMO GPC for a $NU = 1$ case (see Section 4.5.3).

### 4.5.3 Analytical Solution for Constrained MIMO GPC

Constrained GPC for $NU = 1$ in an $m$-input $n$-output case can be described as below:

$$
\min_{\Delta u(t)} J = \left[ \hat{G} \Delta u(t) + f - w \right]^T R \left[ \hat{G} \Delta u(t) + f - w \right] + \Delta u(t)^T Q \Delta u(t)
$$

subject to:

$$
\begin{align*}
\mathbf{u}_{\min} \leq u(t) & \leq \mathbf{u}_{\max} \\
\Delta \mathbf{u}(t)_{\min} \leq \Delta \mathbf{u}(t) & \leq \Delta \mathbf{u}_{\max}
\end{align*}
$$

(4.99)

$$
Y_{\min} = \begin{bmatrix}
\mathbf{y}_{\min} \\
\mathbf{y}_{\min} \\
\vdots \\
\mathbf{y}_{\min}
\end{bmatrix} \leq Y = \begin{bmatrix}
\mathbf{y}(t + 1) \\
\mathbf{y}(t + 2) \\
\vdots \\
\mathbf{y}(t + N2)
\end{bmatrix} \leq \begin{bmatrix}
\mathbf{y}_{\max} \\
\mathbf{y}_{\max} \\
\vdots \\
\mathbf{y}_{\max}
\end{bmatrix} = Y_{\max}
$$

where

$$
\mathbf{u}(t) = [u_1(t) \ u_2(t) \ldots u_m(t)]^T
$$

$$
\Delta \mathbf{u}(t) = [\Delta u_1(t) \ \Delta u_2(t) \ldots \ \Delta u_m(t)]^T
$$

(4.100)

$$
\mathbf{y}(t) = [y_1(t) \ y_2(t) \ldots y_n(t)]
$$
Chapter 4. Constrained Multivariable Model-Based Predictive Control

Obviously, the solution obtained by simply clipping the control signals to their associated bounds is no longer optimal. To obtain an optimal solution, the quadratic programming is used based on the computation of the Kuhn-Tucker multipliers. After substituting $u(t) = \Delta u(t) + u(t-1)$ and $Y = \tilde{G}\Delta u(t) + f$ into Equation 4.99, the inequality constraints are mapped to a linear combination of the control increment $\Delta u(t)$ as follows:

$$u_{\text{min}} \leq u(t) \leq u_{\text{max}}$$

$$u_{\text{min}} - u(t-1) \leq \Delta u(t) \leq u_{\text{max}} - u(t-1)$$

and

$$Y_{\text{min}} \leq Y \leq Y_{\text{max}}$$

$$\tilde{G}^T\tilde{G}^{-1}\tilde{G}^T(Y_{\text{min}} - f) \leq \Delta u(t) \leq \tilde{G}^T\tilde{G}^{-1}\tilde{G}^T(Y_{\text{max}} - f)$$

Combining Equations 4.101 and 4.102 with $\Delta u(t)_{\text{min}} \leq \Delta u(t) \leq \Delta u(t)_{\text{max}}$ gives:

$$a_1 = \Delta u(t) - \alpha \geq 0$$

$$a_2 = -\Delta u(t) + \beta \geq 0$$

where $\alpha$ and $\beta$ are the vectors of the dimension $m$, representing $m$ constraints, and given by:

$$\alpha = \max \{ \Delta u_{\text{min}}, u_{\text{min}} - u(t-1), (\tilde{G}^T\tilde{G})^{-1}\tilde{G}^T(Y_{\text{min}} - f) \}$$

$$\beta = \min \{ \Delta u_{\text{max}}, u_{\text{max}} - u(t-1), (\tilde{G}^T\tilde{G})^{-1}\tilde{G}^T(Y_{\text{max}} - f) \}$$

The vectors $a_1$ and $a_2$ in Equation 4.103 are the linear combination of the constraints, defined by:

$$a_1 = [a_{11} a_{12} \cdots a_{1m}]^T$$

$$a_2 = [a_{21} a_{22} \cdots a_{2m}]^T$$

After substituting Equation 4.103 into the objective function in Equation 4.99, the constrained control problem can then be solved as an unconstrained control problem as:

$$\min_{\Delta u(t)} J = [\tilde{G}\Delta u(t) + f - w]^T R [\tilde{G}\Delta u(t) + f - w]$$
The above equation can be modified as the following form:

$$
\frac{\partial J}{\partial \Delta u(t)} = 2(\tilde{G}^T R \tilde{G} + Q) \Delta u(t) + 2(f - w)^T R \tilde{G} \Delta u(t) + \mu_1^T a_1 + \mu_2^T a_2
$$

(4.109)

The steps involved in the calculation of \( \Delta u(t) \) are based on Mutha's method [99].
Step 1: Determine an unconstrained control solution \( \Delta u^*(t) \).

Step 2: Test to determine if any of the constraints in Equation 4.103 are violated. If all \( a_{ij} \geq 0 \) (no constraints violation) then go to Step 5.

Step 3: Saturate a constraint and verify if the solution satisfies the conditions of Kuhn-Tucker multipliers method, i.e.

- assume any one or two (say \( a_{ik} \) and \( a_{jl} \)) of the constraints are satisfied by the constrained optimal solution, which gives a value \( \Delta u(t) \). Therefore, all other \( \mu \)'s other than \( \mu_{ik} \) and \( \mu_{jl} \) are zero.
- calculate the Kuhn-Tucker multipliers \( \mu_{ik} \) and \( \mu_{jl} \) with the help of Equation 4.109, and verify if the Kuhn-Tucker multipliers \( \mu_{ik} \) and \( \mu_{jl} \) are < 0.

If the Kuhn-Tucker multipliers are negative, then go to Step 5.

Step 4: Repeat Step 3 for the next (set of) constraint(s).

Step 5: Implement the control action and repeat Steps 1-5 at next sample period.

4.6 Constrained Control via Mixed-Weights Least-Squares Algorithm

The constrained control solution in the previous section is obtained by solving a quadratic programming problem. The problem stemmed from a QP-based optimal solution calculation is that the shape of the set of feasible solutions is an irregular volume limited by hyperplanes defined by the constraints [107]. Furthermore, if the problem has no feasible solution, a separate procedure must ensure that a reasonable compromise is achieved between the unsatisfied constraints [107]. The above problems can be avoided by applying an alternative mixed-weights least-squares (MWLS) algorithm. MWLS was proposed by
Rossiter and Kouvaritakes [108] based on a simple modification of the Lawson’s algorithm (the $\infty$-norm minimization).

A least-squares (LS) algorithm is a tool normally used for the approximation of a data set by a finite set of functions using the sum of squares of the errors as a measure of the goodness of the fit. LS has been used in almost every scientific field since it was first developed in the early 19th century by Gauss [109]. Its popularity is due in part to the fact that it can be easily understood and used. As a problem-solving technique, LS minimization is always convenient but not always meaningful or even desirable, because although the sum of the squared deviations may be small, there is no guarantee that one or more of the summands is small [110]. To solve the problem, Lawson [111] proposed a numerical procedure that consists of solving a weighted least-squares problem and recursively modifying the weights such that the larger errors receive higher weights. It was shown in Lawson’s paper [111] that the algorithm always converges to a solution that minimizes the largest error (i.e. the $\infty$-norm). Later, Lawson’s algorithm was analyzed by Rice [112] and extended by Rice and Usow [113] to include the cases $p \in [2, \infty]$. Recently, Lawson’s algorithm was modified by Rossiter and Kouvaritakis [108] to solve control problems with input-output constraints. Because Rossiter and Kouvaritakis method partitions the error vector into two components, one that takes the control objective into account, the other which is for the constraints, this method is referred to as the mixed-weights least-squares algorithm. Correspondingly, the method uses two different recursive relationships for the update of the weights. Very recently, the algorithm of Rice and Usow [113] was extended by Gendron [110, 107] to include the cases $p \in [1, \infty]$.

Addition to QP based optimal solution calculation, the mixed-weights least-squares algorithm is used in the following to solve predictive control problems with input and output constraints. Using MWLS method, a control performance index defined by desired control strategy such as GPC can be easily augmented and the weights can be modified
in a manner that encompasses both the requirement for the future control movements to lie inside the feasible region, and to minimize the control performance index. If the constrained optimization problem is unfeasible, the MWLS will converge to the point that minimizes the 'maximum constraint violation'. Exact theoretical optimality can be obtained only as the number of iterations within the MWLS tends to infinity [108], but the solutions that are optimal within the tolerances of practical implementations can be obtained after a few iterations only. Therefore, there is a significant reduction in the computational burden, especially when compared to the use of QP. A significant advantage of using the MWLS procedure over QP is that by minimizing an appropriate \( \infty \)-norm, it minimizes the likelihood of future constraint violations, and in certain cases overcomes unfeasibility problems altogether. Furthermore, the MWLS application is simple to understand and easy to implement. It can be used to solve constrained control problems for more general cases and with more flexibility.

4.6.1 Preliminary

Linear \( p \)-norm problems

Squared errors in solving an approximation problem is the special case of a \( p \)-norm problem. A \( p \)-norm problem is described as follows:

\[
\| Y \|_p = \begin{cases} 
\left( \sum_{k=1}^{N} y_k^p \right)^{1/p} & \text{if } p \in [1, \infty) \\
\sup_{y_k} |y_k| & \text{if } p = \infty
\end{cases}
\]  

(4.112)

and its distance functional is thus:

\[
\| Y - \hat{Y} \|_p \triangleq \left( \sum_{k=1}^{N} |e_k|^p \right)^{1/p}
\]  

(4.113)

where

\[ e_k = y_k - \hat{y}_k \]  

(4.114)
In Equation 4.113, $\| \cdot \|^p$ denotes the Euclidean $p$-norm. The vector $Y = [y_1 \ y_2 \ \cdots \ y_N]^T$ is the set of observations to be approximated. $\hat{Y} = [\hat{y}_1 \ \hat{y}_2 \ \cdots \ \hat{y}_N]^T$ is the set of approximates to $Y$. If the approximation is considered as linear in parameters, the approximation may be expressed by:

$$\hat{Y}(\theta) = X(\theta) = \sum_{k=1}^{N} x_k^T \theta \quad (4.115)$$

where $\theta$ is a $M$-dimensional vector of adjustable parameters. $X = [x_1 \ x_2 \ \cdots \ x_N]^T$ is a $N \times M$ matrix.

$\|Y - \hat{Y}\|$ in Equation 4.113 is a function that measures the distance between members of $Y$ and members of $\hat{Y}$. The approximation is judged according to the minimization of this measurement. Minimizing the $p$-norm of the distance in Equation 4.113 gives [110]:

$$\nabla_{\theta} = \left( \sum_{k=1}^{N} |e_k|^p \right)^{1/(p-1)} \sum_{k=1}^{N} |e_k|^{p-2} e_k \nabla_{\theta} e_k = 0 \quad (4.116)$$

If assuming $e_k \neq 0$, above equation becomes:

$$\sum_{k=1}^{N} |e_k|^{p-2} e_k \nabla_{\theta} e_k = 0 \quad (4.117)$$

**Least-squares algorithm - a linear 2-norm problem**

A least-squares problem is the special case of $p$-norms with $p = 2$. A least-squares or 2-norm problem can be expressed as follows by setting $p = 2$ in Equation 4.113:

$$\min_{\theta} \|Y - \hat{Y}(\theta)\|^2_2 \quad (4.118)$$

Substituting Equation 4.115 into the above equation and minimizing with respect to $\theta$ gives:

$$\hat{\theta} = (X^T X)^{-1} X^T Y \quad (4.119)$$

or

$$\hat{\theta} = \left( \sum_{k=1}^{N} x_k x_k^T \right)^{-1} \sum_{k=1}^{N} x_k y_k \quad (4.120)$$
Weighted least-squares algorithm - a linear $\infty$-norm problem

Solving a 2-norm problem (Equation 4.118) is to obtain the minimization of the sum of the squared approximation errors. But obviously, it does not guarantee that one or more of the summands is small. To improve this, Lawson [111] proposed weighted least-squares by attaching weights to the error vector in Equation 4.118, i.e., $E = Y - \hat{Y}(\theta)$. The associated weighted least-squares algorithm is described by:

$$\min_{\theta} \| W^{1/2}(Y - \hat{Y}(\theta)) \|_2^2$$

where $W$ is a diagonal matrix with positive real diagonal elements, i.e., $W = \text{diag}\{w_1, w_2, \ldots, w_N\}$.

Computing the gradient of the above equation with respect to $\theta$ and setting it to zero yield:

$$\hat{\theta} = (X^T W X)^{-1} X^T W Y$$  \hspace{1cm} (4.122)

or alternatively

$$\hat{\theta} = \left( \sum_{k=1}^{N} w_k x_k x_k^T \right)^{-1} \sum_{k=1}^{N} w_k x_k y_k$$

Lawson suggested: with $W^{(0)}$ identity matrix, minimizing the objective function of Equation 4.121 for the sequence of weighting matrices $W^{(i)}$ defined recursively by

$$W_{jj}^{(i+1)} = \frac{W_{jj}^{(i)} |e_j^{(i)}|}{\sum_{k=1}^{N} W_{kk}^{(i)} |e_k^{(i)}|}, \quad j = 1 \ldots N$$

where

$$e^{(i)} = y - x^T \theta^{(i)}$$

So, Equation 4.122 can be rewritten as:

$$\hat{\theta} = (X^T W^{(i)} X)^{-1} X^T W^{(i)} Y$$

(4.126)

With increasing $i$, $\hat{\theta}$ will converge to the solution $\theta^*$ [111, 108], which minimizes the largest member of the $Y - \hat{Y}(\theta)$ (hence the $\infty$-norm) because the weights are chosen so
as to place a greater penalty on the member of the error vector which has the largest error.

4.6.2 Constrained Control via Mixed-Weights Least-Squares

Constraints

In GPC based control, input-output constraints may be considered over the control horizon $NU$ and predictive horizon $N_2$. To apply the mixed-weights least squares algorithm, the constraints in the following are mapped in the form of a linear combination of control increments.

**Input amplitude constraints:**

$$u_{\min} \leq u(t) \leq u_{\max}$$
$$u_{\min} \leq u(t+1) \leq u_{\max}$$
$$\vdots$$
$$u_{\min} \leq u(t+NU-1) \leq u_{\max}$$

where $(u_{\min}, u_{\max})$ are the lower and upper bounds on the control amplitudes. The above constraints can be modified in the following form:

$$\frac{|u(t+i-1) - u_{\text{centre}}|}{u_{\text{radius}}} \leq 1$$

$$i = 1 \cdots NU$$

where

$$u_{\text{centre}} = \frac{u_{\max} + u_{\min}}{2} \quad u_{\text{radius}} = \frac{u_{\max} - u_{\min}}{2}$$

After substituting $\triangle u(t+i-1) = u(t+i-1) - u(t+i-2)$ into Equation 4.128, the amplitude constraints on the control actions are mapped in the form of a linear
combination of the future control increments:

\[
\begin{align*}
&\frac{\Delta u(t)}{u_{\text{radius}}} - \frac{u_{\text{centre}} - u(t-1)}{u_{\text{radius}}} \leq 1 \\
&\frac{\Delta u(t) + \Delta u(t+1)}{u_{\text{radius}}} - \frac{u_{\text{centre}} - u(t-1)}{u_{\text{radius}}} \leq 1 \\
&\vdots \\
&\frac{\Delta u(t) + \Delta u(t+1) + \cdots + \Delta u(t + NU - 1)}{u_{\text{radius}}} - \frac{u_{\text{centre}} - u(t-1)}{u_{\text{radius}}} \leq 1
\end{align*}
\] (4.130)

The above equation can be modified in a more compact form as follows:

\[
|L_u \hat{u} - C_u| \leq 1
\] (4.131)

or expressed as an \(\infty\)-norm problem:

\[
||L_u \hat{u} - C_u||_{\infty} \leq 1
\] (4.132)

where \(\hat{u} = [\Delta u(t) \Delta u(t+1) \cdots \Delta u(t + NU - 1)]^T\). \(L_u\) is an \(NU \times NU\) lower diagonal matrix whose non-zero entries are all equal to \(1/u_{\text{radius}}\). \(C_u\) is an \(NU\)-dimensional vector whose entries are all equal to \(u_{\text{centre}} - u(t-1)/u_{\text{radius}}\).

**Input rate constraints:**

\[
|\Delta u(t)| \leq \Delta u
\]

\[
|\Delta u(t+1)| \leq \Delta u
\]

\[
|\Delta u(t + NU - 1)| \leq \Delta u
\] (4.133)

The above equation can be modified as:

\[
\left|\frac{\Delta u(t)}{\Delta u}\right| \leq 1
\]

\[
\left|\frac{\Delta u(t+1)}{\Delta u}\right| \leq 1
\]

\[
\vdots
\]

\[
\left|\frac{\Delta u(t + NU - 1)}{\Delta u}\right| \leq 1
\] (4.134)

The above equation can be easily rewritten in the convenient form as follows:

\[
|L_{du} \hat{u} - 0_{NU \times 1}| \leq 1
\] (4.135)
or expressed as an $\infty$-norm problem:
\begin{equation}
||L_{du} \hat{u} - 0_{NU \times 1}||_\infty \leq 1 \tag{4.136}
\end{equation}
where $L_{du}$ is an $NU \times NU$ diagonal matrix whose non-zero entries are all equal to $1/\Delta u$.

Output amplitude constraints:
\begin{equation}
\begin{bmatrix}
y_{\min} \\
y_{\min} \\
\vdots \\
y_{\min}
\end{bmatrix} \leq \begin{bmatrix}
y(t + 1) \\
y(t + 2) \\
\vdots \\
y(t + N_2)
\end{bmatrix} \leq \begin{bmatrix}
y_{\max} \\
y_{\max} \\
\vdots \\
y_{\max}
\end{bmatrix} \tag{4.137}
\end{equation}

The above output constraints can be modified as:
\begin{equation}
|y(t + j) - y_{\text{centre}}|/\text{yradius} \leq 1 \\
j = 1, \cdots, N_2 \tag{4.138}
\end{equation}
where
\begin{equation}
y_{\text{centre}} = \frac{y_{\max} + y_{\min}}{2} \quad \text{yradius} = \frac{y_{\max} - y_{\min}}{2} \tag{4.139}
\end{equation}

As derived earlier, the output prediction function for a SISO GPC case is expressed by:
\begin{equation}
y = G' \hat{u} + f \tag{4.140}
\end{equation}

To map the output constraints into the form of a linear combination of the future control increments, substituting Equation 4.140 into 4.138 gives:
\begin{equation}
|L_y \hat{u} - C_y| \leq 1 \tag{4.141}
\end{equation}
or expressed as an $\infty$-norm problem:
\begin{equation}
||L_y \hat{u} - C_y||_\infty \leq 1 \tag{4.142}
\end{equation}
where $L_y = G'/\text{yradius}$ is an $N2 \times NU$ matrix. $C_y = (y_c - f)/\text{yradius}$ is an $N_2 \times 1$ matrix. $y_c$ is an $N_2 \times 1$ matrix whose entries are all equal to $y_{\text{centre}}$. 
Combining the constraints in Equations 4.132, 4.136 and 4.142 gives:

\[ \|L\hat{u} - C\|_{\infty} \leq 1 \]  
(4.143)

where

\[
L = \begin{bmatrix}
L_u \\
L_{du} \\
L_y 
\end{bmatrix}, \quad C = \begin{bmatrix}
C_u \\
0 \\
C_y 
\end{bmatrix}
\]  
(4.144)

4.6.3 Control Law

The control problem for SISO GPC subject to the constraints in Equation 4.143 can be re-expressed as:

\[
\min_{\hat{u}} J = [G'\hat{u} + f - w]^T[G'\hat{u} + f - w] + \rho \hat{u}^T\hat{u}
\]

subject to:

\[ \|L\hat{u} - C\|_{\infty} \leq 1 \]

(4.145)

The above control performance index can be modified in the form of a 2-norm minimization as:

\[
\min_{\hat{u}} J = \|R\hat{u} - V\|_2^2
\]

(4.146)

where

\[
R = \begin{bmatrix}
G' \\
\rho^{1/2}I_{NU}
\end{bmatrix}, \quad V = \begin{bmatrix}
w - f \\
0_{NU \times 1}
\end{bmatrix}
\]

(4.147)

\[ R \] is a matrix whose dimensions are \((N_2 + NU) \times NU\) and \(V\) is a vector of the dimension \(N_2 + NU\).
Chapter 4. Constrained Multivariable Model-Based Predictive Control

The minimization of the performance index (Equation 4.146) subject to the constraints (Equation 4.143) is calculated by solving the following mixed-objective (or mixed-weights least-squares problem) function:

\[
J_{MWLS} = \min_{\tilde{u}(i)} \left\| \left( \begin{bmatrix} w^{(i)}I \\ W^{(i)} \end{bmatrix} \right)^{1/2} \left( \begin{bmatrix} R \\ L \\ V \\ C \end{bmatrix} \right) \right\|_2^2
\] (4.148)

where \( w^{(i)} \) is a positive scalar weight associated to the control performance index and \( W^{(i)} \) is a positive definite diagonal matrix associated with the constraints, i.e., \( W^{(i)} = diag \{ w^{(i)}_{11} \ldots w^{(i)}_{jj} \ldots w^{(i)}_{mm} \} \). The weights are updated with:

\[
w^{(i+1)} = \frac{w^{(i)}}{\sum_{k=1}^{m} w^{(i)}_{kk} |e^{(i)}_{k}|} \quad w^{(i+1)}_{jj} = \frac{w^{(i)}_{jj} |e^{(i)}_{j}|}{\sum_{k=1}^{m} w^{(i)}_{kk} |e^{(i)}_{k}|}
\] (4.149)

where \( w^{(i)}_{jj} \) are the entries of \( W^{(i)} \). The initial values of \( w^{(0)} \) and \( W^{(0)} \) are normally set to identity. In the above equation, \( m \) is the total number of constraints (\( m = 2NU + N_2 \)) and \( e^{(i)}_{k} \) are the entries of the vector \( e^{(i)} = [e^{(i)}_{1} e^{(i)}_{2} \ldots e^{(i)}_{m}]^T \) given by:

\[
e^{(i)} = L\tilde{u}^{(i)} - C
\] (4.150)

The control movement \( \tilde{u}^{(i)} \) in the above equation is obtained by minimizing the mixed objective function in Equation 4.148 with respect to \( \tilde{u} \), which is:

\[
\tilde{u}^{(i)} = \left( \begin{bmatrix} R \\ L \end{bmatrix}^T \begin{bmatrix} w^{(i)}I \\ W^{(i)} \end{bmatrix} \begin{bmatrix} R \\ L \end{bmatrix} \right)^{-1} \begin{bmatrix} R \\ L \end{bmatrix}^T \begin{bmatrix} w^{(i)}I \\ W^{(i)} \end{bmatrix} \begin{bmatrix} V \\ C \end{bmatrix}
\] (4.151)

Repeat the procedure with increasing \( i \) until the algorithm converges. A criterion based on the relative variation of the solution from one iteration to the next is calculated to stop the iterations. The criterion is:

\[
s^i = \frac{||\tilde{u}^{(i)} - \tilde{u}^{(i-1)}||_2^2}{||\tilde{u}^{(i)}||_2^2}
\] (4.152)
if \( s^i \begin{cases} > s^* & \text{continue} \\ \leq s^* & \text{stop} \end{cases} \) (4.153)

where \( s^* \) is some threshold.

By recursively using the procedure with increasing \( i \), the algorithm can only converge to the vector \( \hat{u} \) which minimizes the mixed objective function in Equation 4.148 and it always converges to either [108]

\[
\hat{u}_{QP} \quad \text{or} \quad \min_{\hat{u}} \Vert L\hat{u} - C \Vert_{\infty}
\]

(4.154)

depending on whether a feasible solution exists or not. The proof of this convergence and a discussion of its associated properties is given in [108].

Because the algorithm uses two different recursive relationships for the update of the weights, it is referred to as the mixed-weights least-squares algorithm.

### 4.7 Simulations

The purpose of simulations was to demonstrate the effectiveness of the proposed constrained GPC (described earlier) with the application of the quadratic programming and mixed-weights least-squares algorithms. A single-input single-output linear plant was first selected for implementing the control algorithm. A 2 \times 2 linear plant was then selected for further investigation of the proposed control algorithm.

**Example 4.1:** Consider the single-input single-output second-order linear plant:

\[
y(t) - 1.5y(t - 1) + 0.6y(t - 2) = 0.006u(t - 1) + 0.06u(t - 2)
\]

(4.155)

In the following simulations, GPC-based controller tuning parameters were set to \( N_1 = 1, N_2 = 10, NU = 1 \) and \( \rho = 0 \). All the initial parameter estimates were 0.1. The initial covariance matrices for the RLS algorithm were \( 1000 \times I \), and all the forgetting factors
Figure 4.30: Constrained GPC via analytical QP \((t \geq 300\) sampling intervals: \(-1.5 \leq u_t \leq 1.5, -2.0 \leq \Delta u_t \leq 2.0\))
Figure 4.31: Constrained GPC via MWLS (t ≥ 300 sampling intervals: -1.5 ≤ ut ≤ 1.5, -2.0 ≤ Δut ≤ 2.0)
Simulations were performed with and without constraints. The output setpoint was square waves of amplitudes 1 unit.

**Input Amplitude and Rate Constraints:** Figures 4.30 and 4.31 show the behaviors of the constrained control through solving the quadratic programming and mixed-weights least-squares problems, respectively. To make a comparison, the standard unconstrained GPC was simulated for the first 300 sampling intervals. The constrained control algorithm was then turned on at sampling interval 300. The maximum and the minimum amplitude constraints on the control input were set to +1.5 and −1.5. The maximum and the minimum rate constraints on the control input were set to +2.0 and −2.0. As can be seen, the proposed control algorithms based on both quadratic programming and mixed-weights least-squares worked well since the rate and amplitude constraints on the input were successfully limited to their specified ranges without a sacrifice of setpoint tracking.

For the constrained control based on the mixed-weights least-squares algorithm, we set the threshold for the stopping of the algorithm to 0.001 (i.e. \( s^* = 0.001 \) in Equation 2.24). After about 20 iterations, the solution of minimizing the mixed objective function was obtained and the iterations stopped.

**Example 4.2:** Consider the 2 \( \times \) 2 second-order linear plant:

\[
\begin{align*}
y_1(t) &= 1.57y_1(t-1) - 0.615y_1(t-2) + 0.002u_1(t-1) - 0.00213u_1(t-2) + 0.072u_2(t-1) - 0.054u_2(t-2) \\
y_2(t) &= 1.57y_2(t-1) - 0.615y_2(t-2) + 0.0702u_1(t-1) - 0.053u_1(t-2) + 0.03u_2(t-1) - 0.0246u_2(t-2)
\end{align*}
\]

For all the simulations in the following, GPC-based controller tuning parameters were set to \( N_1 = 1, N_2 = 10, NU = 1 \) and \( \rho_1 = \rho_2 = 0 \). Simulations were performed with and without the constraints. The output setpoints were square waves of amplitudes 1 unit.
Input Amplitude and Rate Constrained: Figures 4.33 and 4.34 show the behaviors of the constrained GPC with input amplitude/rate constraints using quadratic programming and mixed-weights least-squares algorithms, respectively.

In Figures 4.33 and 4.34, the standard unconstrained GPC was used for the first 300 sampling intervals. The setpoint tracking of this period simulation was ideal. But some control moves were relatively larger in amplitude and rate. To reduce the amplitude and rate of the control inputs, the proposed constrained GPC was used for the next 100 sampling intervals with \(-4.0 \leq u_1 \leq 4.0, -4.0 \leq \Delta u_1 \leq 4.0, -4.0 \leq u_2 \leq 4.0\) and \(-4.0 \leq \Delta u_2 \leq 4.0\). As can be seen in Figures 4.33 and 4.34, the input amplitude and rate were limited to their specified ranges for the sampling interval \(\geq 300\) without a sacrifice of setpoint tracking.

For the constrained control based on MWLS, we set the threshold for the stopping of the algorithm to 0.001 (i.e. \(s^* = 0.001\) in Equation 2.24). After about 50 iterations, the solution of minimizing the mixed objective function was obtained and the iterations stopped. Figure 4.32 shows the feasible value of \(s^*\) as a function of the iteration number.
Output Constraints: Figures 4.35 and 4.36 show the behaviors of the output constrained GPC using quadratic programming and mixed-weights least-squares algorithm, respectively. The standard unconstrained GPC was used for $t < 300$ sampling intervals and the setpoint tracking of the simulations was ideal. To reduce the small overshoots on the first output $y_1$, the constrained GPC was on at $t = 300$ sampling interval and $y_1$ was limited to $-1.0 \leq y_1 \leq 1.0$. Figures 4.35 shows that the constrained GPC via quadratic programming was successful since the output $y_1$ was limited to its desired range. To further investigate the constrained control algorithm by solving the MWLS, the maximum and minimum amplitudes on both $y_1$ and $y_2$ were set to 1.0 and $-1.0$. Figure 4.36 shows the time-trajectory plot of the simulation of the constrained GPC using the mixed weights least-squares algorithm.
Figure 4.33: Constrained GPC via analytical QP (\( t \geq 300 \) sampling intervals: \(-4.0 \leq u_1 \leq 4.0, -4.0 \leq u_2 \leq 4.0, |\Delta u_1| \leq 4.0, |\Delta u_2| \leq 4.0\))
Figure 4.34: Constrained GPC via MWLS (t ≥ 300 sampling intervals: -4.0 ≤ u₁ ≤ 4.0, -4.0 ≤ u₂ ≤ 4.0, |Δ u₁| ≤ 4.0, |Δ u₂| ≤ 4.0)
Figure 4.35: Output constrained GPC via analytical QP (t ≥ 300 sampling intervals: -1.0 ≤ y₁ ≤ 1.0)
Figure 4.36: Output constrained GPC via MWLS (t ≥ 300 sampling intervals: -1.0 ≤ y₁ ≤ 1.0, -1.0 ≤ y₂ ≤ 1.0)
4.8 Summary and Conclusions

An analytical solution exists for the unconstrained generalized predictive control (GPC). However, no analytical solution exists for the general case of constrained GPC. The solution for the general case of constrained GPC may be obtained by using commercial optimization software packages, but it turns out to be computationally demanding as compared with an analytical solution. This chapter describes the derivation of a constrained MIMO GPC subject to input and/or output constraints. An analytical solution of the constrained MIMO GPC for the case of \( NU = 1 \) was obtained through solving a quadratic programming (QP) problem by computing Kuhn Tucker multipliers. A problem resulting from a QP-based optimal solution calculation is that there will be no feasible solutions if one or more constraints are overly stringent. To avoid the problem, the mixed-weights least-squares (MWLS) algorithm was introduced to the constrained MIMO GPC to solve an optimization problem. A significant advantage of using the MWLS is that if the constrained optimization problem is unfeasible, the MWLS will converge to the point that minimizes the maximum constraint violation. In addition, the MWLS-based optimal solution is able to handle the general cases of constrained GPC with more flexibility. The proposed control schemes were tested on SISO and \( 2 \times 2 \) linear models, showing the success of handling input and/or output constraints. The proposed control schemes will be used for the control of the TMP process as given in Chapter 6.
5.1 Introduction

In dealing with a model-based control approach, parametric representations, e.g. transfer functions such as an ARMAX or ARIMA model, are commonly used. A possible reason for this scheme is that the model is simple and the theory based on this model is well developed [114]. However, a linear transfer function-based control approach will work well only when the nonlinearity of a plant is not severe and the actual plant is represented properly by the structure of the model. In the case of severe nonlinearity, the linear approximation-based control may not be satisfactory. It may result in, in some cases, significant static errors and in other cases, oscillation or even instability [114]. Also, the linear transfer function-based approaches generally require accurate knowledge of the model order and time-delay, and can be very sensitive to a model mismatch. Alternatively, a non-parametric representation like the impulse response requires no other information than the linearity of the plant [115]. A disadvantage of the impulse response representation is that a large number of parameters are needed.

Considered in this chapter is the use of the orthonormal functions or filter representations to present the dynamics of an actual plant. By the right choice of an orthonormal filter, plant dynamics can be represented properly without the need of assumptions about plant order and time delay. An interesting set of orthonormal filters is the set of Laguerre functions. To overcome the problems caused by severe nonlinearities, the use of nonlinear
Laguerre functions representations is the nature choice. The nonlinear Laguerre function is a Wiener type model, which is composed of a linear dynamic block followed by a memoryless nonlinear function. The nonlinear Laguerre function approximation-based control will be very beneficial for the control of wood chip refining when the nonlinearity between refiner motor load and plate gap becomes severe, i.e., the incremental gain between refiner motor load and plate gap subject to a sudden change in the sign (Figure 6.50). This is particularly critical when a refiner is operating close to the maximum load point. The linear approximation-based control for this case may lead to unstable operation. Alternatively, the nonlinear Laguerre functions are used to approximate process dynamics and static nonlinearity. The Laguerre function is expressed in a state-space form, so any state-space control design techniques may be used for a state-feedback control. However, the generalized predictive control algorithm is used for the controller development because of its simplicity and ease of use. The optimal control actions are calculated by minimizing future output errors and control movements. The future outputs are predicted based on the Laguerre model under the influence of future control actions. For an adaptive control implementation, the nonlinear Laguerre model can be arranged in a linear form in the parameters so that the powerful but simple Recursive Least Squares (RLS) algorithm can be used for the parameter estimation.

The use of orthonormal functions for obtaining approximations goes back as far as the development of the Fourier series [116]. However, an appropriate and well-known orthogonal basis set is the set of Laguerre filters [117]. The studies and applications of Laguerre approximations have been carried out recently by many researchers. Linear Laguerre model-based adaptive control theory was first proposed by Dumont and Zervos [118]. Their theory was then applied to pH control by Dumont, et al. [119], and industrial application results showed that the Laguerre model-based adaptive pH control was
superior to those using other types of models [119]. Laguerre filter-based nonlinear dynamic approximation is based on Volterra functional series expansion [120, 114]. Early work on the Volterra model directly applies the nonlinear memoryless function to the plant input signal and its delays, resulting in a very large number of model parameters even for a simple nonlinear process [114]. However, using Laguerre filter-based nonlinear approximation the number of model parameters can be significantly reduced [120]. In a paper by Fu and Dumont [51], a Laguerre filter-based nonlinear dynamic model was used for the adaptive control of the motor load of a wood chip refiner used in the pulp and paper industry. Fu and Dumont's work was further extended by Du et al. [52] into a $2 \times 2$ control system based on mixed linear and nonlinear Laguerre functions for the control of both the motor load and the outlet consistency. Other studies on the theories and applications of the Laguerre approximation can be found in [116, 121, 114]. Studies and discussions of time domain properties of Laguerre approximation are detailed in [121]. Laguerre function-based representation is often used because of its flexible structure so that the model complexity can easily be changed on-line. Also, Laguerre function-based representation is very similar to transient signals due to its block-oriented feature and thus can be understood easily.

First, linear and nonlinear Laguerre function-based approximations are introduced in Section 5.2. A linear Laguerre function-based GPC algorithm is derived in Section 5.3. In the same section, the limitation of a linear model-based control is described and a nonlinear Laguerre function-based control is then presented. Section 5.4 presents the technique for the estimation of Laguerre function parameters. Summary and conclusions are given in Section 5.5.
5.2 The Laguerre Functions

The Laguerre functions, a complete orthonormal set in \( L_2[0, \infty) \), have often been used because of their convenient network realizations [117] and their similarity to transient signals [122]. In the following, a linear Laguerre model is a linear representation where the model output is a linear combination of the Laguerre filter outputs, while a nonlinear Laguerre model is a nonlinear representation where the model output is a nonlinear combination of the Laguerre filter outputs.

5.2.1 The Linear Laguerre Functions

The continuous-time Laguerre functions

The continuous-time linear Laguerre model has been discussed and used for the control purposes in [116, 119, 121]. In the \( s \)-domain, the \( i \)th order Laguerre function \( F_i(s) \) is as follows:

\[
F_i(s) = \sqrt{2p} \frac{(s - p)^{i-1}}{(s + p)^i}
\]

(5.157)

Any transfer function of a stable system can be described by this Laguerre ladder network using infinite number of filters as:

\[
G(s) = \sum_{i=1}^{\infty} c_i F_i(s)
\]

(5.158)

In practice, the first \( N \) filters are used to approximate the system. This approximation can be adjusted by the convergence results given in [121]. \( p \) is the Laguerre filter pole. The output of the \( N \)th order Laguerre model is:

\[
y(t) = \sum_{i=1}^{N} c_i l_i(t)
\]

(5.159)

where \( l_i(t) \) is the \( i \)th Laguerre function output. The Laguerre ladder network is also illustrated in Figure 5.37(a) where the first block is a low-pass filter and the remaining
blocks are all-pass filters. As can be seen from Figure 5.37(a), the simplicity of a phase-shift chain, convenient network realization and the similarity to transient signals are the significant advantages of the Laguerre function representations.

The discrete-time Laguerre functions

The transformation of the continuous-time Laguerre functions to the discrete ones has been discussed in [122, 123]. A discrete-time Laguerre ladder network, similar to the continuous one, is illustrated in Figure 5.37(b). In the z-domain, the \( i \)th order Laguerre filter is:

\[
F_i(z) = \sqrt{1-a^2} \frac{(1-az)^{i-1}}{(z-a)^i}
\]  (5.160)

Any stable transfer function \( G(z) \) can be expanded as:

\[
G(z) = \sum_{i=1}^{\infty} c_i F_i(z)
\]  (5.161)
If the first $N$ filters are used to approximate a process, the output of the $N$th order discrete-time Laguerre model is:

$$y(t) = \sum_{i=1}^{N} c_i l_i(t)$$

(5.162)

where $l_i(t)$ is the $i$th Laguerre filter output and $c_i$ are the Laguerre function coefficients.

The state space formula of the discrete-time Laguerre ladder network can be easily derived from Figure 5.37(b) as:

$$L(t) = AL(t - 1) + Bu(t - 1)$$

$$y(t) = C^T L(t)$$

(5.163)

where

$$C = [c_1, c_2, \cdots, c_N]^T$$

$$L(t) = [l_1(t), l_2(t), \cdots, l_N(t)]^T$$

(5.164)

and

$$A = \begin{bmatrix}
    a_i & 0 & 0 & \cdots & 0 \\
    1 - a_i^2 & a_i & 0 & \cdots & 0 \\
    (-a_i)(1 - a_i^2) & 1 - a_i^2 & a_i & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    (-a_i)^{N-2}(1 - a_i^2) & (-a_i)^{N-3}(1 - a_i^2) & \cdots & \cdots & a_i
\end{bmatrix}$$

(5.165)

$$B = \sqrt{1 - a_i^2} \begin{bmatrix}
    1 \\
    (-a_i) \\
    (-a_i)^2 \\
    \vdots \\
    (-a_i)^{N-1}
\end{bmatrix}$$

As can be seen from Equation 5.165, the matrix $A$ and the vector $B$ depend on the Laguerre filter pole $a$ which is a free parameter. The choice of a Laguerre filter pole was first discussed in [124] and more recently in [125].
5.2.2 The Nonlinear Laguerre Functions

The nonlinear Laguerre model is a special case of the general Wiener-Volterra series representation [126], where the Wiener-Volterra functional series is truncated and the Volterra kernels are approximated using the Laguerre functions [120] with the assumption that all the kernels are in $L_2[0, \infty)$.

**The continuous-time nonlinear Laguerre functions**

The Wiener-Volterra series representation for a nonlinear and time invariant system is:

$$y(t) = h_0 + \sum_{n=1}^{\infty} \int_0^\infty \cdots \int_0^\infty h_n(\tau_1, \cdots, \tau_n)u(t - \tau_1) \cdots u(t - \tau_n) d\tau_1 \cdots d\tau_n$$  \hspace{0.5cm} (5.166)

where $u(t)$ is the input and $y(t)$ is the output. The functions $h_n(\tau_1 \cdots \tau_n)$ are the Volterra kernels which represent the nonlinear dynamic of the system. In practice, the system is often approximated by a finite Wiener-Volterra series. To simplify the notation, the above series is truncated to have only the first two Volterra kernels $h_1(\tau_1)$ and $h_2(\tau_1, \tau_2)$. Thus we have:

$$y(t) = h_0 + \int_0^\infty h_1(\tau_1)u(t - \tau_1)d\tau_1 + \int_0^\infty \int_0^\infty h_2(\tau_1, \tau_2)u(t - \tau_1)u(t - \tau_2)d\tau_1d\tau_2$$  \hspace{0.5cm} (5.167)

Let $\phi_i(t)$ be the $i$th-order Laguerre function. The $i$th order Laguerre filter output is:

$$l_i(t) = \int_0^\infty \phi_i(\tau)u(t - \tau)d\tau$$  \hspace{0.5cm} (5.168)

Since all the Volterra kernels form a complete orthonormal set in $L_2[0, \infty)$, the truncated Volterra kernels using $N$ Laguerre filters are expressed as [120]:

$$h_1(\tau_1) = \sum_{k=1}^{N} c_k \phi_k(\tau_1) \hspace{0.5cm} h_2(\tau_1, \tau_2) = \sum_{n=1}^{N} \sum_{m=1}^{N} c_{nm}\phi_n(\tau_1)\phi_m(\tau_2)$$  \hspace{0.5cm} (5.169)

where $c_k$ and $c_{nm}$ are constant coefficients. Substituting the above equation into Equation 5.167 and considering the orthonormality of the Laguerre functions give:

$$y(t) = c_0 + \sum_{k=1}^{N} c_k l_k(t) + \sum_{n=1}^{N} \sum_{m=1}^{N} c_{nm} l_n(t)l_m(t)$$  \hspace{0.5cm} (5.170)
Chapter 5. Laguerre Function-Based Adaptive-Predictive Control

The continuous-time state-space formula of the 2nd order nonlinear Laguerre model is then expressed as:

\[
\dot{L}(t) = AL(t) + Bu(t)
\]

\[
y(t) = c_0 + C^T L(t) + L^T(t)DL(t)
\]

where

\[
C = [c_1, \cdots, c_N]^T
\]

\[
L(t) = [l_1(t), \cdots, l_N(t)]^T
\]

\[
D = \begin{bmatrix}
c_{11} & c_{12} & \cdots & c_{1N} \\
c_{21} & c_{22} & \cdots & c_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
c_{N1} & c_{N2} & \cdots & c_{NN}
\end{bmatrix}
\]

Since the Volterra kernels are symmetric, i.e., \(c_{mn} = c_{nm}\), the number of unknown coefficients in Equation 5.170 is \((N+1)(N+2)/2\). The matrix \(A\) and the vector \(B\) in Equation 5.171 depend only on the selected Laguerre filter pole.

**The discrete-time nonlinear Laguerre functions**

Similarly to the continuous-time nonlinear Laguerre formula, the discrete-time nonlinear Laguerre model is expressed as follows by extending the continuous-time formula using the discrete-time Laguerre filters:

\[
L(t) = AL(t-1) + Bu(t-1)
\]

\[
y(t) = c_0 + C^T L(t) + L^T(t)DL(t) + \cdots
\]

where the scalar \(c_0\), the vector \(C\) and the matrix \(D\) are the Laguerre function coefficients to be estimated. Under the condition of neglecting the higher-order Volterra kernels, the above equation is then an approximation of Wiener-Volterra series representation for a nonlinear dynamic system. The discrete-time nonlinear Laguerre network is also illustrated in Figure 5.38, where the linear dynamic part, in occurrence as the Laguerre
network, is followed by a memoryless nonlinear function which is a nonlinear combination of the Laguerre filter outputs.

5.3 SISO Laguerre Function-Based Predictive Control

5.3.1 Linear Model-Based Control

Model output prediction

Consider a stochastic process described by the following linear Laguerre model:

\[
L(t) = AL(t-1) + Bu(t-d)
\]

\[
y(t) = C^T L(t) + \xi(t)/\Delta
\]

where \(u(t)\) is the input and \(y(t)\) is the output. \(\xi(t)\) is uncorrelated, normally distributed zero-mean random noise. The difference operator \(\Delta = 1 - z^{-1}\) in the denominator of the noise term forces an integrator into the controller in order to eliminate any offset between the measured output and its setpoint. The matrix \(A\) and the vector \(B\) are dependent on
only the selected Laguerre filter poles (Equation 5.165).

If Equation 5.174 is multiplied by \( \Delta \) with the assumption that the time-delay \( d = 1 \) (without losing the generality by adding states) and the future control movements \( \Delta u(t + j) = 0 \), then we have:

\[
\Delta L(t + j) = A^j \Delta L(t) + (A^{j-1} + \cdots + I)B \Delta u(t)
\]

To derive a \( j \)-step ahead predictor \( \hat{y}(t + j) \), consider the following identity:

\[
y(t + j) = y(t) + \Delta y(t + 1) + \cdots + \Delta y(t + j)
\]

\[
y(t) + \sum_{i=1}^{j} \Delta y(t + i)
\]

Multiplying Equation 5.175 by \( \Delta z^j \) and combining with Equations 5.176 and 5.177 gives:

\[
y(t + j) = y(t) + \sum_{i=1}^{j} C^T A^i \Delta L(t) + \sum_{i=1}^{j} C^T (A^{i-1} + \cdots + I)B \Delta u(t) + \sum_{i=1}^{j} \xi(t + i)
\]

As the noise components in the above equation are all in the future, the output predictor, given measured output and control action up to time \( t \), is:

\[
\hat{y}(t + j) = \sum_{i=1}^{j} C^T (A^{i-1} + \cdots + I)B \Delta u(t) + \sum_{i=1}^{j} C^T A^i \Delta L(t) + y(t)
\]

The output predictor \( \hat{y}(t + j) \) consists of two parts: the first term depending on the control action yet to be determined, and the rest depending on the past control action and the measured output up to time \( t \).

Considering the output predictor \( \hat{y}(t + j) \) over the predictive horizon \( N_2 \) gives:

\[
\hat{y}(t + 1) = C^T A^0 B \Delta u(t) + C^T A \Delta L(t) + y(t)
\]

\[
\hat{y}(t + 2) = \sum_{i=1}^{2} C^T (A^{i-1} + \cdots + I)B \Delta u(t) + \sum_{i=1}^{2} C^T A^i \Delta L(t) + y(t)
\]

\[
\vdots
\]

\[
\hat{y}(t + N_2) = \sum_{i=1}^{N_2} C^T (A^{i-1} + \cdots + I)B \Delta u(t) + \sum_{i=1}^{N_2} C^T A^i \Delta L(t) + y(t)
\]
The above equation can be rewritten in a more compact form as:

\[ \hat{y} = G \triangle u(t) + f \quad (5.181) \]

where

\[
G = \begin{bmatrix}
C^T A^0 B \\
\sum_{i=1}^{2} C^T (A^{i-1} + \cdots + I)B \\
\vdots \\
\sum_{i=1}^{N_2} C^T (A^{i-1} + \cdots + I)B
\end{bmatrix}
\quad (5.182)
\]

and

\[
f = \begin{bmatrix}
C^T A \triangle L(t) + y(t) \\
\sum_{i=1}^{2} C^T A^i \triangle L(t) + y(t) \\
\vdots \\
\sum_{i=1}^{N_2} C^T A^i \triangle L(t) + y(t)
\end{bmatrix}
\quad (5.183)
\]

Predictive control law

The target of GPC-based predictive control is to minimize the following performance index:

\[
J = E \left\{ \sum_{j=N_1}^{N_2} [y(t + j) - w(t + j)]^2 + \sum_{j=1}^{NU} \rho \triangle u(t + j - 1)^2 \right\}
\quad (5.184)
\]

The above performance index can be rewritten in a more compact form:

\[
J = E \left\{ (y - w)^T (y - w) + \rho \tilde{u}^T \tilde{u} \right\}
\quad (5.185)
\]

where

\[
y = [y(t + N_1), y(t + N_1 + 1), \cdots, y(t + N_2)]^T
\]
\[
w = [w(t + N_1), w(t + N_1 + 1), \cdots, w(t + N_2)]^T
\quad (5.186)
\]
\[
\tilde{u} = [\triangle u(t), \triangle u(t + 2), \cdots, \triangle u(t + NU - 1)]^T
\]

Substituting Equation 5.181 into Equation 5.185 for the case \( N_1 = 1, NU = 1 \) and minimizing \( J \) yields:

\[
\triangle u(t) = (G^T G + \rho)^{-1} G^T (w - f)
\quad (5.187)
\]
Estimation of Laguerre model parameters

The output of the $N$-filter Laguerre network in Equation 5.175 is simply the linear combination of all filter outputs, i.e.

$$y(t) = \sum_{i=1}^{N} c_i l_i(t) + \xi(t)/\triangle$$

(5.188)

The above equation can then be rearranged in the form:

$$\triangle y(t) = \theta_i^T \phi_t + \xi(t)$$

(5.189)

where the model parameters to be estimated are:

$$\theta_i^T = [c_1 \ c_2 \ \cdots \ c_N]$$

(5.190)

and the regression vector is:

$$\phi_t^T = [\triangle l_1(t) \ \triangle l_2(t) \ \cdots \ \triangle l_N(t)]$$

(5.191)

Since Equation 5.189 is a linear form in the parameters, the standard RLS parameter estimation algorithm can be used for the adaptive control purpose.

Limitations of linear model-based control

Control of a wood chip refiner in a TMP process is sometimes difficult due to the nonlinear relationship between the refiner motor load and plate gap with a sudden change in the sign of incremental gain, illustrated by the dashed line in Figure 5.39. For the nonlinearity represented by the solid curve in Figure 5.39, the linear model 1 based control will be satisfactory as long as the output setpoint $y_r$ is reachable. For this case, the process will settle at the operating point $A$. However, when the nonlinearity is represented by the dashed line, the linear model 2 based on the operating point $B$ will predict that point $C$ is the desired position. The control action will move the process to the position $D$. The linear model based control will then predict that the desired position
Figure 5.39: Linear model-based control will lead to oscillation when the nonlinearity is represented by the dashed curve [1]
is $F$, but resulted control action leads the process back to $B$ point. If no heuristic is used, the process will be in a limited cycle and can never be stabilized at the extreme point $E$ because at that point the corresponding process gain is zero, resulting in an infinite control signal. One solution to the problem is the use of a nonlinear model-based control approach.

5.3.2 Nonlinear Model-Based Control

Model output prediction

Consider a 2nd order nonlinear Laguerre model using $N$ filters:

$$L(t) = AL(t - 1) + Bu(t - d)$$
$$y(t) = c_0 + C^T L(t) + L^T DL(t)$$

Recursive use of Equation 5.192 gives:

$$L(t + j) = A^j L(t) + (A^{j-1} + \cdots + I)Bu(t)$$
$$= A^j L(t) + \bar{A}_j Bu(t)$$
Chapter 5. Laguerre Function-Based Adaptive-Predictive Control

where

\[ \tilde{A}_j = A^{j-1} + \cdots + I \]  

(5.194)

For simplifying the above derivation, we assumed that the time delay is one (i.e. \( d = 1 \)) and the future control movements are zeros (i.e. \( u(t) = u(t+1) = \cdots = u(t+j) \)).

**Predictive control law**

Predictive control action is obtained by minimizing the performance index:

\[ J = \sum_{j=N_1}^{N_2} [w(t+j) - y(t+j)]^2 + \rho \triangle u^2(t) \]  

(5.195)

The minimization of the above performance index is done by solving:

\[ \frac{\partial J}{\partial u(t)} = -2 \sum_{j=N_1}^{N_2} [w(t+j) - y(t+j)] \frac{\partial y(t+j)}{\partial u(t)} + 2\rho \triangle u(t) = 0 \]  

(5.196)

or

\[ \rho \triangle u(t) = \sum_{j=N_1}^{N_2} [w(t+j) - y(t+j)] \frac{\partial y(t+j)}{\partial u(t)} \]  

(5.197)

Substituting the output equation in Equation 5.192 into the above equation gives:

\[ \rho \triangle u(t) = \sum_{j=N_1}^{N_2} [w(t+j) - y(t+j)] [C^T + 2L^T(t+j)D] \tilde{A}_j B \]  

(5.198)

The above equation will involve solving a nonlinear equation and the analytical solution for a general case is difficult to obtain. An analytical solution for the special case of \( N_1 = N_2 = d \) and \( \rho = 0 \) is derived in [51] which is:

\[ [a \ u^2(t) + b \ u(t) + c][2 \ a \ u(t) + b] = 0 \]  

(5.199)

The control solution is then:

\[ u(t) = \begin{cases} \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} & \text{if } b^2 - 4ac \geq 0 \\ -\frac{b}{2a} & \text{if } b^2 - 4ac < 0 \end{cases} \]  

(5.200)
where
\[ a = B^T \bar{A}_d^T D \bar{A}_d B \]
\[ b = C^T \bar{A}_d B + 2L^T(t)(A^d)^T D \bar{A}_d B \] (5.201)
\[ c = c_0 + C^T A^d L(t) + L^T(t)(A^d)^T D A^d L(t) - w(t) \]

and \( w(t) = w(t + j), j \geq 1 \), i.e., the future setpoints are considered to be equal to the present one.

If it is the case of \( b^2 - 4ac > 0 \), one way to select the sign in Equation 5.200 is to look at how the future steps will be influenced, as suggested in [127]. However, if it is the case of \( b^2 - 4ac < 0 \), the solution in Equation 5.200 is corresponds to the extreme point in Figure 5.39.

**Estimation of Laguerre model parameters**

The Laguerre model output given in Equation 5.192 can be rearranged as:

\[
y(t) = c_0 + c_1 l_1(t) + \cdots + c_N l_N(t) + c_{11} l_1^2(t) + c_{12} l_1(t) l_2(t) + \cdots \\
+ c_{1N} l_1(t) l_N(t) + c_{22} l_2^2(t) + \cdots + c_{NN} l_N^2(t) \] (5.202)

\[
\theta^T \phi_t
\]

where the vector of the model parameters to be estimated is:

\[
\theta_t^T = [c_0 \ c_1 \ \cdots \ c_N \ c_{11} \ c_{12} \ \cdots \ c_{22} \ \cdots \ c_{NN}] \] (5.203)

and the regression vector is:

\[
\phi_t^T = [1 \ l_1 \ \cdots \ l_N \ l_1^2 \ l_1 l_2 \ \cdots \ l_2^2 \ \cdots \ l_N^2] \] (5.204)

As can be seen, the nonlinear Laguerre model output is simply in a linear form in the model parameters. Therefore, the standard RLS parameter estimation algorithm can be used for the model parameter estimation.
5.4 MIMO Laguerre Function-Based Predictive Control

In the following, we use refiner control as an example to describe a multivariable Laguerre function-based predictive control.

5.4.1 Mixed Linear-Nonlinear Laguerre Model

Consider a $2 \times 2$ mixed linear-nonlinear Laguerre functions using the first $N$ filters for modelling a wood chip refiner:

\[
\begin{align*}
L_{1,t} &= A_1 L_{1,t-1} + B_1 y_{1,t-d_1} \\
L_{2,t} &= A_2 L_{2,t-1} + B_2 y_{2,t-d_2} \\
y_{1,t} &= c_0 + C_1^T L_{1,t} + L_{1,t}^T D_1 L_{1,t} + \cdots + k_1 y_{2,t-1} + N_{1,t} \\
y_{2,t} &= C_2^T L_{2,t} + k_2 y_{1,t-1} + N_{2,t}
\end{align*}
\]

where $u_1$ is the pulse sent to the hydraulic cylinder to position the gap, $u_2$ is the dilution flow rate to control the consistency in the refining zone, $y_1$ is the motor load and $y_2$ the outlet consistency. $d_1$ and $d_2$ present time delays in terms of sampling unit for each input-output pair. $N_1$ and $N_2$ are the unmeasured disturbance and measurement noises acted on the outputs. $k_1 y_2$ and $k_2 y_1$ present the load disturbances induced due to the changes in the motor load and outlet consistency. Matrix $A_i$ and vector $B_i$ depend on
only the Laguerre filter poles $a_i$:

$$A = \begin{bmatrix}
    a_i & 0 & 0 & \cdots & 0 \\
    1 - a_i^2 & a_i & 0 & \cdots & 0 \\
    (-a_i)(1 - a_i^2) & 1 - a_i^2 & a_i & \cdots & 0 \\
    \vdots & \cdots & \cdots & \cdots & \cdots \\
    (-a_i)^{N-2}(1 - a_i^2) & (-a_i)^{N-3}(1 - a_i^2) & \cdots & \cdots & a_i
\end{bmatrix}$$

(5.206)

$$B = \sqrt{1 - a_i^2} \begin{bmatrix}
    1 \\
    (-a_i) \\
    (-a_i)^2 \\
    \vdots \\
    (-a_i)^{N-1}
\end{bmatrix}$$

and

$$L_1 = [l_{11} \ l_{12} \ \cdots \ l_{1N}]^\top \quad C_1 = [c_{11} \ c_{12} \ \cdots \ c_{1N}]^\top$$

$$L_2 = [l_{21} \ l_{22} \ \cdots \ l_{2N}]^\top \quad C_2 = [c_{21} \ c_{22} \ \cdots \ c_{2N}]^\top$$

(5.207)

$$D_1 = \begin{bmatrix}
    d_{11} & d_{12} & \cdots & d_{1N} \\
    d_{21} & d_{22} & \cdots & d_{2N} \\
    \vdots & \vdots & \vdots & \vdots \\
    d_{N1} & d_{N2} & \cdots & d_{NN}
\end{bmatrix}$$

(5.208)

The discussions on the selection of Laguerre poles was first given in [124] and recently in [114].

The Laguerre model outputs in Equation 5.205 can be rearranged in a linear in parameter form:

$$y_{1,t} = \theta_{1,t} \phi_{1,t}$$

$$y_{2,t} = \theta_{2,t} \phi_{2,t}$$

(5.209)
Chapter 5. Laguerre Function-Based Adaptive-Predictive Control

where

\[ \theta_1 = [c_0, c_{11}, c_{12}, d_{11}, d_{12}, d_{22}, k_1] \]

\[ \phi_{1,t} = [1, l_{11}, l_{12}, l_{11}^2, l_{11} l_{12}, l_{12}^2, y_{2,t-1}]^T \] (5.210)

and

\[ \theta_2 = [c_{21}, c_{22}, k_2] \]

\[ \phi_{2,t} = [l_{21}, l_{22}, y_{1,t-1}]^T \] (5.211)

By the above arrangement, the simple but powerful RLS is readily implemented. This estimation algorithm is:

\[ \hat{\theta}_{i,t} = \hat{\theta}_{i,t-1} + P_{i,t}\phi_{i,t}^T e_{i,t} \]

\[ P_{i,t} = \lambda_i^{-1}[I - \frac{P_{i,t-1}\phi_{i,t}\phi_{i,t}^T}{\lambda_i + \phi_{i,t}^T P_{i,t-1}\phi_{i,t}}]P_{i,t-1} \] (5.212)

\[ e_{i,t} = y_{i,t} - \phi_{i,t}^T \hat{\theta}_{i,t-1} \]

where \( P_{i,t} \) are the error covariance matrix, \( e_{i,t} \) are the predictive error and \( \lambda_i \) are the forgetting factors.

5.4.2 MIMO Laguerre Function-Based Predictive Control Law

The future control actions are determined by minimizing a designed trade-off between control performance and constraints on the control movements as:

\[ J = E \left\{ \sum_{j=-N_1}^{N_2} \left[ (y_{1,t+j} - w_{1,t+j})^2 + (y_{2,t+j} - w_{2,t+j})^2 \right] \right\} + \rho_1 \Delta u_{1,t}^2 + \rho_2 \Delta u_{2,t}^2 \] (5.213)

where \( w_i \) are the setpoints of the outputs and \( \Delta u_{i,t} \) are the control increments (\( \Delta u_{i,t} = u_{i,t} - u_{i,t-1} \)).

Assuming time-delay \( d_i = 1 \) (without losing the generality by adding states) and
recursively substituting Equation 5.205 itself yields:

\[
L_{1,t+j} = A_1^jL_{1,t} + (A_1^{j-1} + \cdots + I)B_1u_{1,t}
\]

\[
= A_1^jL_{1,t} + \bar{A}_1B_1u_{1,t} \tag{5.214}
\]

\[
L_{2,t+j} = A_2^jL_{2,t} + (A_2^{j-1} + \cdots + I)B_2u_{2,t}
\]

\[
= A_2^jL_{2,t} + \bar{A}_2B_2u_{2,t}
\]

It was assumed in the above equation that the future control movements \( \triangle u_{1,t+j} = 0, \ j \geq 0 \) (or \( u(t) = u(t + 1) = \cdots = u(t + j) \)). To derive a \( j \)-step ahead predictor of \( y_{1,t+j} \), the output functions in Equation 5.205 are rewritten as:

\[
y_{1,t+j} = c_0 + C_1^T L_{1,t+j} + L_{1,t+j}^T D_1 L_{1,t+j} + \cdots + k_1 y_{2,t+j-1} + \xi_{1,t+j}
\]

\[
y_{2,t+j} = C_2^T L_{2,t+j} + k_2 y_{1,t+j-1} + \xi_{2,t+j}/\triangle
\]

where \( \xi_{1,t} \) are the independent, normally distributed (white) noise sequence with variance \( \rho^2_\xi \). The noise term \( N_{1,t} \) acted on \( y_{1,t} \) in Equation 5.205 is assumed to be white noise to simplify the derivation.

To derive a \( j \) step ahead predictor of \( y_{2,t+j} \), consider the following identify:

\[
y_{2,t+j} = y_{2,t} + \triangle y_{2,t+1} + \cdots + \triangle y_{2,t+j}
\]

\[
= y_{2,t} + \sum_{i=1}^{j} \triangle y_{2,t+i} \tag{5.216}
\]

Substituting Equation 5.214 into Equation 5.215 and combining the resultant with Equation 5.216 gives:

\[
y_{1,t+j} = c_0 + C_1^T [A_1^jL_{1,t} + \bar{A}_1B_1u_{1,t}] + \\
+ [A_1^jL_{1,t} + \bar{A}_1B_1u_{1,t}]^T D_1 [A_1^jL_{1,t} + \bar{A}_1B_1u_{1,t}] + k_1 y_{2,t+j-1} \tag{5.217}
\]

\[
y_{2,t+j} = y_{2,t} + \sum_{i=1}^{j} C_2^T [A_2^j \triangle L_{2,t} + \bar{A}_2B_2 \triangle u_{2,t}] + \\
+ \sum_{i=1}^{j} k_2 \triangle y_{1,t+i-1}
\]
The optimum control actions can then be determined by substituting the above equation into the performance index in Equation 5.213 and minimizing it, i.e.:

\[
\frac{\partial J}{\partial u_{1,t}} = 2 \sum_{j=N_1}^{N_2} [y_{1,t+j} - w_{1,t+j}] [C_1^T + 2L_{1,t+j}^T D_1] \Delta_1 B_1 + 2\rho_1 \Delta u_{1,t} = 0
\]

\[
\frac{\partial J}{\partial u_{2,t}} = 2 \sum_{j=N_1}^{N_2} (y_{2,t+j} - w_{2,t+j}) \sum_{i=1}^{j} C_2^T (A_2^{i-1} + \cdots + I) B_2 + 2\rho_2 \Delta u_{2,t} = 0
\]

Let \( N_1 = N_2 = d \), the above equation can then be simplified as:

\[
\rho_1 \Delta u_{1,t} = (a u_{1,t}^2 + b u_{1,t} + c)(2 a u_{1,t} + b)
\]

\[
\rho_2 \Delta u_{2,t} = \gamma_1 \Delta u_{2,t} + \gamma_0
\]

where

\[
a = B_1^T \bar{A}_1^T D_1 \bar{A}_1 B_1
\]

\[
b = C_1^T \bar{A}_1 B_1 + 2L_{1,t}^T (A_1^d)^T D_1 \bar{A}_1 B_1
\]

\[
c = c_0 + C_1^T A_1^d L_{1,t} + L_{1,t}^T (A_1^d)^T D_1 A_1^d L_{1,t} + k_1 y_{2,t+d-1} - w_1
\]

\[
\gamma_1 = -[\sum_{i=1}^{d} C_2^T (A_2^{i-1} + \cdots + I) B_2]^2
\]

\[
\gamma_0 = [w_2 - y_{2,t} - \sum_{i=1}^{d} C_2^T A_2^i \Delta L_{2,t} - \sum_{i=1}^{d} k_2 \Delta y_{1,t+i-1}] \sum_{i=1}^{d} C_2^T (A_2^{i-1} + \cdots + I) B_2
\]

If \( \rho_1 \) is set to zero, i.e., \( \rho_1 = 0 \), the above equation can be solved analytically as follows:

\[
u_{1,t} = \begin{cases} 
  -\frac{b \pm \sqrt{b^2 - 4ac}}{2a} & \text{if } b^2 - 4ac \geq 0 \\
  -\frac{b}{2a} & \text{if } b^2 - 4ac < 0
\end{cases}
\]

\[
u_{2,t} = \frac{\gamma_0}{\rho_2 - \gamma_1} + u_{2,t-1}
\]

In the case of \( b^2 - 4ac \geq 0 \), one way to select the sign in Equation 5.221 is to look at how the future steps will be influenced, as suggested by Wittenmark [127]. However, for the control of the refiner load, minus sign in the above equation is used since it corresponds to the plate position in the normal operating region [51]. When \( b^2 - 4ac < 0 \), the only solution is \( u_{1,t} = -\frac{b}{2a} \) which corresponds to the extreme point.
5.5 Summary and Conclusions

A linear transfer function-based control will work well only when the nonlinearity of a plant is not severe and the model structure error is not significant. In the case of severe nonlinearity and unknown dynamics of a process, the linear transfer function based control may lead to remarkable static errors or unstable operations. To avoid the problems, the nonlinear Laguerre functions—a type of orthonormal functions are used to represent an actual plant. By the right choice of an orthonormal filter, the plant can be represented properly without the need for assumptions about plant order and time delay, i.e., accurate assumptions about these true values are not necessary. The Laguerre function-based approximation is very similar to transient signals due to its block-oriented feature and thus can be understood easily. The Laguerre functions are expressed in a state-space form, so any state-space control design techniques may be used for a state-feedback controller. However, the GPC algorithm is used for the controller design because of its simplicity and ease of use. The optimal control actions are calculated by minimizing future output errors and control movements. The future outputs are predicted based on a mixed linear-nonlinear Laguerre model under the influence of future control actions. For an adaptive control application, the Laguerre model can be arranged in a linear form in model parameters so that the powerful but simple RLS algorithm can be used for the parameter estimation. The proposed Laguerre model-based control approach, an alternative to a linear transfer function-based control scheme, will be used to handle the severe nonlinearity of a TMP process, i.e., the incremental gain between refiner motor load and plate gap is subject to sudden changes in the sign. The is particularly critical when a refiner is operating close to the maximum load point (see Chapter 6).
Chapter 6

Constrained Multivariable Control Of A TMP Plant

6.1 Introduction

A thermo-mechanical pulping (TMP) process is a rather complex process. Good control will improve the production, efficiency and pulp quality, while poor control can lead to operating difficulties and poor physical and papermaking properties. The development of a control strategy suitable for the process is a challenging task due to the multivariable nature, variables' interactions, nonlinearity, time varying, time delay and stochastic disturbances of the process. For the reasons of equipment physical limitations, stable operation and product quality requirement, the plant has to be operated under a number of constraints, e.g. the maximum load attainable by each refiner, the maximum hydraulic closing pressure of a hydraulic loading system, and the desired ranges for the dilution flow and the refining consistency. The controller for this particular process must thus be able to handle the problems efficiently together with the robustness to modelling errors.

In the past, various control strategies were proposed towards a control of the process and the product quality. However, many of them were based on the quality-energy relationship, i.e., by adjusting specific energy alone to control pulp quality such as freeness. This is not satisfactory. The earlier research work shows that at least two property variables should be used to predict pulp properties rather than just using freeness alone [66]. Also, a given specific energy may produce fibres with substantially different characteristics. Figure 6.40 shows that the given specific energy, $E_1 = E_2$, may produce the fibres
with substantially different characteristics, where $E_1$ results in fibre cutting and $E_2$ leads to fibre fibrilization. Actually, pulp quality is mainly a function of two variables: the number of refining bar impacts $N$ per unit of pulp and the specific energy per bar impact or refining intensity $I$. The right combination of these two variables will completely determine the pulp quality. Another drawback associated with the past control approaches is that, in many cases, the control calculations were made and implemented under the assumption that the resulted control actions were able to be implemented. This is not satisfactory because the inability to implement the control signal exists with the concerns of equipment physical limitations, large variations on variables and product quality requirement. Implementing control actions without considering process constraints may lead to degraded process performance and even closed-loop instability. The limitations associated with the past control methods motivated this study towards a better control of a TMP plant.

This chapter presents a novel control strategy for the control of a TMP plant by applying the control theories that have been developed and presented in the thesis. The control objective is to achieve specified pulp quality by selecting correct setpoints of the
operating variables: specific energy, refining intensity, production rate, motor load and refining consistency. The setpoints of the operating variables are achieved by automatically manipulating the closing pressure of a hydraulic loading system, the chip screw speed and the dilution flow rate, without violation of process constraints.

The chapter is structured as follows. Section 6.2 describes a choice of variables and the proposed control strategy for the two-stage TMP plant. Section 6.3 presents the control strategy for the control of TMP refiners and simulation results. Laguerre function-based control strategy is developed in Section 6.4 to overcome severe nonlinearity of the process. Summary and conclusions are given in Section 6.5.

6.2 Choice of Variables and Overall Control Strategy

6.2.1 Choice of Variables

As mentioned earlier, many pulp properties in terms of papermaking like freeness, fibre length distribution, fibre specific surface, shive content, coarseness, flexibility and so on can be used to describe pulp quality. However, most researchers agree that at least two property variables should be used to define pulp quality. Under this study, the following pulp properties:

- Freeness, $CSF$
- Long fiber content, $LF$
- Shive content, $SC$

are chosen to predict pulp quality. The choice of these variables is because (i) these variables can adequately describe the handsheet, strength and drainage properties of mechanical pulp [70], and (ii) they can be reasonably measured on-line at the present
Chapter 6. Constrained Multivariable Control Of A TMP Plant

The targets of pulp properties, normally given by plant operators or from higher level plant optimization systems, are generally expressed using a quality window as:

<table>
<thead>
<tr>
<th>Pulp Properties</th>
<th>Lower Bound</th>
<th>Desired Value</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeness (ml)</td>
<td>$CSF_{\min}$</td>
<td>$\leq CSF^* \leq$</td>
<td>$CSF_{\max}$</td>
</tr>
<tr>
<td>Long fibre content (%)</td>
<td>$LF_{\min}$</td>
<td>$\leq LF^* \leq$</td>
<td>$LF_{\max}$</td>
</tr>
<tr>
<td>Shive content (%)</td>
<td>$SC_{\min}$</td>
<td>$\leq SC^* \leq$</td>
<td>$SC_{\max}$</td>
</tr>
</tbody>
</table>

The subscripts min, max refer to the lower and upper bounds in the pulp properties. For a given raw material and refiner design, the pulp properties are strongly affected by the following operating variables at each stage:

- Production rate, $F_c$
- Specific energy, $E$
- Refining intensity, $I$ or specific refining power, $\dot{e}$
- Motor load, $P$
- Refining consistency, $C$

To achieve the setpoints of the above operating variables according to pulp quality requirement, the following input variables are manipulated at each stage:

- Closing pressure of a hydraulic load system to position the gap, $P_c$
- Chip transfer screw speed to achieve the desired production rate, $S_c$
- Dilution water flow rate to control the refining consistency, $F_d$
6.2.2 Constraints

A number of constraints need to be considered during automatic operation of a TMP plant.

Amplitude Constraints. The maximum load attainable and the maximum hydraulic pressure are bounded to avoid plate clashes and for the reason of product quality requirement. The outlet consistency is limited with concerns of both pulp quality requirement and stable operation of a refiner because operation instability may occur in a condition of high refining consistency. The dilution flow rate at each stage is also bounded to avoid valve saturation and a plate clash because plate clashes may happen when the dilution flow rate is increased up to a certain level [73]. The described possible amplitude constraints are summarized as follows:

<table>
<thead>
<tr>
<th>Amplitude Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Motor Load</td>
</tr>
<tr>
<td>Consistency</td>
</tr>
<tr>
<td>Closing Pressure</td>
</tr>
<tr>
<td>Dilution flow</td>
</tr>
</tbody>
</table>

Rate Constraints. For the reasons of large variations on the variables and physical limitations associated with actuators or instrument, the rate constraints in the closing pressure, dilution flow rate and chip screw speed are considered as follows:

<table>
<thead>
<tr>
<th>Rate Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Closing Pressure</td>
</tr>
<tr>
<td>Dilution flow</td>
</tr>
<tr>
<td>Chip Screw Speed</td>
</tr>
</tbody>
</table>
A controller for this particular process should not be allowed to drive the process variables outside the specified rate and amplitude constraints by taking those constraints into account in the controller design.

### 6.2.3 Control Strategy Over Two Stages

As can be seen from Equation 2.40, pulp quality is expressed as a function of the specific energy $E$, refining intensity $I$ and specific refining power $\dot{e}$. However, specific refining power $\dot{e}$ can be determined from refining intensity $I$ as given in Equation 2.36. Therefore,

![Block diagram for the control of a two-stage TMP plant](image-url)

**Figure 6.41:** Block diagram for the control of a two-stage TMP plant
pulp quality is mainly a function of specific energy $E$ and refining intensity $I$. The right combination of $E$ and $I$ will completely determine the pulp quality for a given raw material.

Figure 6.41 illustrates the proposed control strategy for the two-stage TMP plant i.e., the desired pulp quality is achieved by selecting the setpoints of $E$ and $I$ for each stage. If there were on-line sensors for $E$ and $I$, there would have automatic control of $E$ and $I$ according to pulp quality requirement. As $E$ is a function of production rate $F_c$ and motor load $P$ (see Equation 2.23), and $I$ is function of refining consistency $C$ for given specific energy and disc rotational speed (see Equations 2.31, 2.32 and 2.33), $E$ and $I$ for each stage are controlled through selecting the setpoints of production rate which is set by chip the screw speed $S_c$, motor load $P$ and outlet consistency $C$. The manipulated variables at each stage are the chip screw speed $S_c$ (for the 1st stage), closing pressure $P_c$ and dilution flow rate $F_d$. The details regarding the controller design for each stage are given in Section 6.3.

To calculate the total specific energy $E_t$ and refining intensity $I_t$ for the two-stage refining, the following equations can be used:

$$E_t = E_1 + E_2$$
$$I_t = E_t / N_t$$

where

$$N_t = N_1 + N_2$$

The subscripts 1, 2 refer to the first and second stage refining respectively, and the subscript, $t$, refers to the total stage refining. $N$ is the number of refiner bar impact received by a unit pulp (Equation 2.33).
6.3 Control of TMP Refiners

For control purposes, a refiner is considered as a unit operation with three inputs; (i) the wet mass flow rate of chips/pulp, (ii) the hydraulic closing pressure forcing the plates together and (iii) the dilution flow rate to control refining zone consistency. The outputs of a refiner are the specific energy, pulp production rate, motor load, outlet consistency, refining intensity and pulp quality. The process is highly coupled, as most of the inputs simultaneously affect most outputs. The motor load can be easily measured. Consistency measurements become available with the help of some consistency sensor developed lately. Pulp property measurements can be obtained using quality sensors available on the market. In many mills, the production rate is predicted from chip screw speed. Because of the lack of associated sensors, the specific energy and the refining intensity will be predicted from the production rate, motor load and refining consistency based on mechanistic models.

The controller for a refiner is developed to achieve the desired specific energy and refining intensity through control of the production rate, motor load and outlet consistency by manipulating the closing pressure, chip screw speed and dilution flow rate.

6.3.1 Process Model

Any model-based control technique depends upon the existence of a mathematical model to represent the input-output relationship. The dynamic relationships between refiner inputs and outputs about some operating point have been identified as first order plus time delay as follows:

\[
P(t) = \frac{\omega_{10}}{1 - \delta_{11} z^{-1}} P_c(t - d_1) + \frac{\omega_{20}}{1 - \delta_{21} z^{-1}} S_c(t - d_2) + N_1(t) \\
C(t) = \frac{\omega_{30}}{1 - \delta_{31} z^{-1}} S_c(t - d_3) + \frac{\omega_{40}}{1 - \delta_{41} z^{-1}} F_d(t - d_4) + d(z^{-1}) P(t - 1) + N_2(t)
\]  

(6.225)
where $d_i$ are the time delay associated with each input-output pair. The load disturbance $d(z^{-1})P(t - 1)$ from the motor load to the consistency is also included to enhance the control performance instead of relying on the load disturbance properties of the controller. $N_i(t)$ represent the joint effect of all unobserved disturbances acting within the process and may be modelled by:

$$N_i(t) = \frac{1 - \theta_i z^{-1}}{\Delta} \xi_i(t) \quad (6.226)$$

where $\xi_i$ are independent, normally distributed (white) noise sequences with variance $\sigma^2_{\xi_i}$.

The dynamic-stochastic model in Equation 6.225 may not be appropriate for online parameter estimation, but it can be placed in the form of a CARIMA model by multiplying both sides by the common denominator polynomial, i.e.

$$A_1(z^{-1})y_1,t = B_1(z^{-1})u_{1,t-d_1} + B_2(z^{-1})u_{2,t-d_2} + \frac{C_1(z^{-1})}{\Delta} \xi_{1,t} \quad (6.227)$$

$$A_2(z^{-1})y_2,t = B_3(z^{-1})u_{2,t-d_3} + B_4(z^{-1})u_{3,t-d_4} + D(z^{-1})y_{1,t-1} + \frac{C_2(z^{-1})}{\Delta} \xi_{2,t}$$

where

$$A_i(z^{-1}) = 1 - a_{i1} z^{-1} - a_{i2} z^{-2}$$

$$B_j(z^{-1}) = b_{j0} + b_{j1} z^{-1}$$

$$C_i(z^{-1}) = 1 - c_{i1} z^{-1} - c_{i2} z^{-2} - c_{i3} z^{-3}$$

$$D(z^{-1}) = d_0 + d_1 z^{-1} + \cdots$$

$$i = 1, 2$$

$$j = 1, \ldots, 4$$

(6.228)

To simplify the notation in Equation 6.227, the inputs and outputs are re-defined as:

$$y_1 : \text{the motor load, } P$$

$$y_2 : \text{the outlet consistency, } C$$

$$u_1 : \text{the hydraulic closing pressure, } P_c$$

$$u_2 : \text{the chip transfer screw speed, } S_c$$

$$u_3 : \text{the dilution flow rate, } F_d$$

(6.229)
Chapter 6. Constrained Multivariable Control Of A TMP Plant

Since many mills use the chip transfer screw to set the desired production rate, the actual screw speed \( u_2 \) is composed of two terms: the target screw speed \( u_2^* \) associated with the desired production rate and its offset \( \nabla u_2 \) manipulated by the controller to compensate for output deviations, i.e.
\[
    u_{2,t} = \nabla u_{2,t} + u_{2,t}^*
\]  
(6.230)

Substituting Equation 6.230 into 6.227 yields:
\[
    A_1(z^{-1})y_{1,t} = B_1(z^{-1})u_{1,t-1} + B_2(z^{-1}) \nabla u_{2,t-1} + B_2(z^{-1})u_{2,t-1}^* + C_1(z^{-1})\xi_{1,t}/\triangle 
\]
\[
    A_2(z^{-1})y_{2,t} = B_3(z^{-1}) \nabla u_{2,t-1} + B_4(z^{-1})u_{3,t-1} + B_3(z^{-1})u_{2,t-1}^* + D(z^{-1})y_{1,t-1} + C_2(z^{-1})\xi_{2,t}/\triangle 
\]  
(6.231)

Note that \( d_i - 1 \) leading zeros have been added to the \( B_i(z^{-1}) \) polynomial in the above equation to simplify the notation, i.e., \( \deg(B_i) = nb_i + d_i - 1 \).

To derive a \( j \)-step ahead predictor \( y_{i,t+j} \) based on Equation 6.231, consider the Diophantine identity:
\[
    C_i(z^{-1}) = E_{i,j}(z^{-1})A_i(z^{-1}) \triangle + z^{-j}F_{i,j}(z^{-1})
\]  
(6.232)

where \( E_{i,j}(z^{-1}) (\deg(E_{i,j}) = j - 1) \) and \( F_{i,j}(z^{-1}) (\deg(F_{i,j}) = \max(na_i, nc_i - j)) \) are determined by given \( C_i(z^{-1}), A_i(z^{-1}) \) and \( j \). The \( C_i(z^{-1}) \) polynomials used for the controller design might differ from the noise terms identified from the process identification stage if the nature of process disturbances or setpoint changes expected in the future is different from that obtained during identification [128].

Multiplying Equation 6.231 by \( E_{i,j} \triangle z^j \) and then combining with Equation 6.232 give:
\[
    y_{1,t+j} = E_{1,j}B_1 \triangle u_{1,t+j-1}^f + E_{1,j}B_2^f \nabla u_{2,t+j-1}^f + E_{1,j}B_2 \triangle u_{2,t+j-1}^* 
\]
\begin{align*}
+ F_{1,j}y_{1,t}^f + E_{1,j}\xi_{1,t+j} \\
y_{2,t+j} &= E_{2,j}B'_3 \triangledown u_{2,t+j-1}^f + E_{2,j}B_4 \Delta u_{2,t+j-1}^f + E_{2,j}B_3 \Delta u_{2,t+j-1}^f + E_{2,j}D \Delta y_{1,t+j-1} + F_{2,j}y_{2,t}^f + E_{2,j}\xi_{2,t+j}
\end{align*}

In the above equation, \( B'_i(z^{-1}) = B_i(z^{-1}) \) effectively eliminates integral action from \( u_{2,t} \). This allows the chip screw speed to be operated about the target production rate. Superscript ' denotes filtering by \( 1/C_i(z^{-1}) \). After eliminating future unknown noise terms \( E_{i,j}\xi_{i,t+j} \), the output predictors can be modified in the form:

\begin{align*}
\hat{y}_{1,t+j} &= G_{1,j} \Delta u_{1,t+j-1}^f + G_{2,j} \triangledown u_{2,t+j-1}^f + G_{2,j}'' \Delta u_{2,t+j-1}^f + F_{1,j}y_{1,t}^f \\
\hat{y}_{2,t+j} &= G_{3,j} \Delta u_{3,t+j-1}^f + G_{4,j} \triangledown u_{2,t+j-1}^f + G_{3,j}'' \Delta u_{2,t+j-1}^f + F_{2,j}y_{2,t}^f + E_{2,j}D \Delta y_{1,t+j-1}
\end{align*}

where

\begin{align*}
G_{1,j} &= E_{1,j}B_1 \\
G_{2,j} &= E_{1,j}B'_2 \\
G_{3,j} &= E_{2,j}B'_3 \\
G_{4,j} &= E_{2,j}B_4 \\
G_{2,j}'' &= E_{1,j}B_2 \\
G_{3,j}'' &= E_{2,j}B_3
\end{align*}

To separate the past known filtered control actions from the current and future un-filtered control actions yet to be determined, consider the identity:

\[ G_{k,j}(z^{-1}) = G'_{k,j}(z^{-1})C_i(z^{-1}) + z^{-j}\Gamma_{k,j}(z^{-1}) \]

(6.236)

where \( G'_{k,j}(z^{-1})(\deg(G_{k,j}(z^{-1}) = j - 1) \) and \( \Gamma_{k,j}(z^{-1})(\deg(\Gamma_{k,j}(z^{-1}) = \max(nb_k - 1, nc_i - 1)) \) are defined by \( G_{k,j}(z^{-1}), C_i(z^{-1}) \) and \( j \). Substituting Equation 6.236 into 6.234 gives:

\begin{align*}
\hat{y}_{1,t+j} &= G'_{1,j} \Delta u_{1,t+j-1}^f + G'_{2,j} \triangledown u_{2,t+j-1}^f + \Gamma_{1,j} \Delta u_{1,t-1}^f + \Gamma_{2,j} \triangledown u_{2,t-1}^f
\end{align*}
Chapter 6. Constrained Multivariable Control Of A TMP Plant

\[ + G''_{2,j} \Delta u_{2,t+j-1} + F_{1,j} y_{1,t} \]  \hspace{1cm} (6.237)

\[ \hat{y}_{2,t+j} = G'_{3,j} \triangle u_{2,t+j-1} + G'_{4,j} \Delta u_{3,t+j-1} + \Gamma_{3,j} \nabla u_{2,t-1} + \Gamma_{4,j} \Delta u_{3,t-1} \]
\[ + G''_{3,j} \Delta u_{2,t+j-1} + E_{2,j} D \triangle y_{1,t+j-1} + F_{2,j} y_{2,t} \]

The first two terms on the right side of Equation 6.237 include the effect of the future control inputs yet to be determined: future closing pressure changes (\(\Delta u_{1,t+j-1}\)), future chip screw speed deviations from the target production rate (\(\nabla u_{2,t+j-1}\)) and future dilution flow rate changes (\(\Delta u_{3,t+j-1}\)). The rest comprises process free response (open-loop predictions), including the past known controls, measured outputs, load disturbance and future changes in the production rate. Since the \(G''_{i,j}\) polynomials in Equation 6.237 are in the order of \(j - 1 + nb_t\), i.e.

\[ G''_{2,j}(z^{-1}) = g''_{2,0} + g''_{2,1} z^{-1} + \ldots + g''_{2,j-1} z^{-(j-1)} + \ldots \]  \hspace{1cm} (6.238)
\[ G''_{3,j}(z^{-1}) = g''_{3,0} + g''_{3,1} z^{-1} + \ldots + g''_{3,j-1} z^{-(j-1)} + \ldots \]

all terms except for the \(j\)-th-term in \(G''_{i,j}(z^{-1}) \triangle u_{2,t+j-1}^f\) involve the past and future changes in the production rate. The production rate can be assumed constant over the predictive horizon if its future changes are unknown. After eliminating the past and future changes in the production rate, Equation 6.237 is then written as:

\[ \hat{y}_{1,t+j} = G'_{1,j} \triangle u_{1,t+j-1} + G'_{2,j} \nabla u_{2,t+j-1} + \Gamma_{1,j} \nabla u_{1,t-1} + \Gamma_{2,j} \nabla u_{2,t-1} \]
\[ + g''_{2,j-1} \triangle u_{2,t}^f + F_{1,j} y_{1,t} \]  \hspace{1cm} (6.239)
\[ \hat{y}_{2,t+j} = G'_{3,j} \nabla u_{2,t+j-1} + G'_{4,j} \Delta u_{3,t+j-1} + \Gamma_{3,j} \nabla u_{2,t-1} + \Gamma_{4,j} \Delta u_{3,t-1} \]
\[ + g''_{3,j-1} \triangle u_{2,t}^f + E_{2,j} D \nabla y_{1,t+j} + F_{2,j} y_{2,t} \]

The above equation can be modified in a more compact form over the predictive and control horizons:

\[ \hat{y}_1 = G'_1 \tilde{u}_1 + G'_2 \tilde{u}_2 + f_1 \]  \hspace{1cm} (6.240)
\[ \hat{y}_2 = G'_2 \tilde{u}_2 + G'_4 \tilde{u}_3 + f_2 \]
where
\[
\hat{y}_1 = [y_{1,t+1} \, y_{1,t+2} \cdots y_{1,t+N_3}]^T
\]
\[
\hat{y}_2 = [y_{2,t+1} \, y_{2,t+2} \cdots y_{2,t+N_3}]^T
\]
\[
\hat{u}_1 = [\Delta u_{1,t} \, \Delta u_{1,t+1} \cdots \Delta u_{1,t+NU-1}]^T
\]
\[
\hat{u}_2 = [\nabla u_{2,t} \, \nabla u_{2,t+1} \cdots \nabla u_{2,t+NU-1}]^T
\]
\[
\hat{u}_3 = [\Delta u_{3,t} \, \Delta u_{3,t+1} \cdots \Delta u_{3,t+NU-1}]^T
\]

and
\[
f_1 = \begin{bmatrix}
\Gamma_{1,1} \Delta u_{1,t-1} + \Gamma_{1,1} \nabla u_{2,t-1} + F_{1,1} y_{1,t} + g_{2,0}^{'''} \Delta u_{2,t} \\
\Gamma_{1,2} \Delta u_{1,t-1} + \Gamma_{1,2} \nabla u_{2,t-1} + F_{1,2} y_{1,t} + g_{2,1}^{'''} \Delta u_{2,t} \\
\vdots \\
\Gamma_{1,N_2} \Delta u_{1,t-1} + \Gamma_{2,N_2} \nabla u_{2,t-1} + F_{1,N_2} y_{1,t} + g_{2,N_2-1}^{'''} \Delta u_{2,t}
\end{bmatrix}
\] (6.242)
\[
f_2 = \begin{bmatrix}
\Gamma_{3,1} \nabla u_{2,t-1} + \Gamma_{4,1} \Delta u_{3,t-1} + F_{2,1} y_{2,t} + g_{3,0}^{'''} \Delta u_{2,t} \\
\Gamma_{3,2} \nabla u_{2,t-1} + \Gamma_{4,2} \Delta u_{3,t-1} + F_{2,2} y_{2,t} + g_{3,1}^{'''} \Delta u_{2,t} \\
\vdots \\
\Gamma_{3,N_2} \nabla u_{2,t-1} + \Gamma_{4,N_2} \Delta u_{3,t-1} + F_{2,N_2} y_{2,t} + g_{3,N_2-1}^{'''} \Delta u_{2,t}
\end{bmatrix}
\] (6.243)

### 6.3.2 Controller Development

Simply stated, the control objective is to maintain the production rate, the motor load and the outlet consistency at their respective setpoints according to the desired specific energy and refining intensity through manipulating the chip screw speed, the closing pressure and the dilution flow rate. Since the number of inputs, in this application, is one greater than the number of outputs, one of the inputs must be constrained about some steady-state value [128]. Many mills use the chip screw to set the desired production rate and, therefore the controller should constrain the chip screw speed about the target production rate while still leaving some room for manipulation. The described control
strategy can be achieved by designing a controller to minimize the following quadratic cost function:

\[
\min_{\Delta u_1, \Delta u_2, \Delta u_3} J = E \left\{ \sum_{j=N_1}^{N_2} \left[ q_1 (y_{1,t+j} - w_{1,t+j})^2 + q_2 (y_{2,t+j} - w_{2,t+j})^2 \right] + \sum_{j=1}^{NU} \left[ \rho_1 \Delta u_{1,t+j-1}^2 + \rho_2 \Delta u_{2,t+j-1}^2 + \rho_3 \Delta u_{3,t+j-1}^2 \right] \right\}
\]

(6.244)

where \( q_i \) are the output weighting factors, \( \rho_i \) are the control weight factors and \( w_i \) are the output setpoints. \( \nabla u_2 \) is the difference between the actual chip meter speed \( u_2 \) and the target production rate \( u_2^* \) \( (u_{2,t} = u_2^* + \nabla u_2) \).

The above quadratic cost function can be modified and put in a more compact form:

\[
\min_{\bar{u}_1, \bar{u}_2, \bar{u}_3} J = q_1 (y_1 - w_1)^T(y_1 - w_1) + q_2 (y_2 - w_2)^T(y_2 - w_2) + \rho_1 \bar{u}_1^T\bar{u}_1 + \rho_2 \bar{u}_2^T\bar{u}_2 + \rho_3 \bar{u}_3^T\bar{u}_3
\]

(6.245)

where \( w_i \) are the vectors of the future setpoints, i.e., \( w_i = [w_{i,t+1} \ w_{i,t+2} \cdots w_{i,t+NU}]^T \). Substituting Equation 6.240 into the above equation gives:

\[
\min_{\bar{u}_1, \bar{u}_2, \bar{u}_3} J = q_1 [G'_1 \bar{u}_1 + G'_2 \bar{u}_2 + f_1 - w_1]^T[G'_1 \bar{u}_1 + G'_2 \bar{u}_2 + f_1 - w_1] + q_2 [G'_3 \bar{u}_3 + G'_4 \bar{u}_4 + f_2 - w_2]^T[G'_3 \bar{u}_3 + G'_4 \bar{u}_4 + f_2 - w_2] + \rho_1 \bar{u}_1^T\bar{u}_1 + \rho_2 \bar{u}_2^T\bar{u}_2 + \rho_3 \bar{u}_3^T\bar{u}_3
\]

(6.246)

As stated earlier, an industrial refiner needs to be operated under a number of constraints. The necessary constraints given earlier are rewritten as:

\[
\begin{align*}
& u_{1\min} \leq u_{1,t+i-1} \leq u_{1\max} \quad |\Delta u_{1,t+i-1}| \leq \Delta u_{1} \\
& u_{3\min} \leq u_{3,t+i-1} \leq u_{3\max} \quad |\Delta u_{3,t+i-1}| \leq \Delta u_{3} \\
& y_{1\min} \leq y_{1,t+j} \leq y_{1\max} \\
& y_{2\min} \leq y_{2,t+j} \leq y_{2\max} \\
& i = 1, \cdots, NU \\
& j = N_1, \cdots, N_2
\end{align*}
\]

(6.247)
If none of the constraints is violated, the control actions \( u^* = [\bar{u}_1 \bar{u}_2 \bar{u}_3]^T \) derived by minimizing Equation 6.246 are optimal. If, however, one or more constraints are violated, \( u^* \) will no longer be optimal. The optimal solution should be derived by minimizing the constrained quadratic cost function in Equation 6.246 with the satisfaction of the constraints in Equation 6.247. The following is the derivation of the optimal solution through solving a quadratic programming (QP) problem and a mixed-weight least-squares (MWLS) problem.

**Optimal solution via analytical QP**

The cost function in Equation 6.246 subject to the constraints in Equation 6.247 can be modified and put in the form of:

\[
\begin{align*}
\min_u J &= \bar{u}^T H \bar{u} + 2c^T \bar{u} \\
\text{subject to:} \quad A(\bar{u}) &\geq 0
\end{align*}
\]  

(6.248)
where

\[ \mathbf{u} = [\mathbf{u}_1^T \mathbf{u}_2^T \mathbf{u}_3]^T \]  

(6.249)

\[ H = \begin{bmatrix}
 h_{11} & h_{12} & 0 \\
 h_{21} & h_{22} & h_{23} \\
 0 & h_{32} & h_{33}
\end{bmatrix} \]  

(6.250)

\[ = \begin{bmatrix}
 q_1 \mathbf{G}_1^T \mathbf{G}_1' + \rho_1 & q_1 \mathbf{G}_1^T \mathbf{G}_2' & 0 \\
 q_1 \mathbf{G}_2^T \mathbf{G}_1' & q_1 \mathbf{G}_2^T \mathbf{G}_2' + q_2 \mathbf{G}_3^T \mathbf{G}_3' + \rho_2 & q_2 \mathbf{G}_4^T \mathbf{G}_4' \\
 0 & q_2 \mathbf{G}_4^T \mathbf{G}_3' & q_2 \mathbf{G}_4^T \mathbf{G}_4' + \rho_3
\end{bmatrix} \]

and

\[ c = \begin{bmatrix}
 c_1 \\
 c_2 \\
 c_3
\end{bmatrix} = \begin{bmatrix}
 q_1 \mathbf{G}_1^T (\mathbf{f}_1 - \mathbf{w}_1) \\
 q_1 \mathbf{G}_2^T (\mathbf{f}_1 - \mathbf{w}_1) + q_2 \mathbf{G}_3^T (\mathbf{f}_2 - \mathbf{w}_2) \\
 q_2 \mathbf{G}_4^T (\mathbf{f}_2 - \mathbf{w}_2)
\end{bmatrix} \]  

(6.251)

Matrix \( A(\mathbf{u}) = [a_1, a_2 \cdots a_6]^T \) is a linear combination of the constraints and the control actions yet to be determined.

Generally, the quadratic cost function in Equation 6.244 solves both for the current input and \( NU - 1 \) future inputs. However, here we use \( NU = 1 \) for all three inputs since this tends to simplify the computations. For the case of \( NU = 1 \), the constraints in Equation 6.247 can be modified in the form (see Appendix B.1.):

\[ \begin{align*}
a_1 &= \triangle u_{1,t} - \alpha_1 \geq 0 \\
a_2 &= -\triangle u_{1,t} + \beta_1 \geq 0 \\
a_3 &= \nabla u_{2,t} - \alpha_2 \geq 0 \\
a_4 &= -\nabla u_{2,t} + \beta_2 \geq 0 \\
a_5 &= \triangle u_{3,t} - \alpha_3 \geq 0 \\
a_6 &= -\triangle u_{3,t} + \beta_3 \geq 0
\end{align*} \]  

(6.252)
where
\[ \alpha_1 = \max(-\Delta u_1, u_{1\text{min}} - u_{1,t-1}, [1 0 0] \cdot P \cdot h_L) \]
\[ \beta_1 = \min(\Delta u_1, u_{1\text{max}} - u_{1,t-1}, [1 0 0] \cdot P \cdot h_H) \]
\[ \alpha_2 = \max(-\Delta u_2 - \Delta u_{2,t}^*, \nabla u_{2,t-1}, [0 1 0] \cdot P \cdot h_L) \]
\[ \beta_2 = \min(\Delta u_2 - \Delta u_{2,t}^* + \nabla u_{2,t-1}, [0 1 0] \cdot P \cdot h_H) \]
\[ \alpha_3 = \max(-\Delta u_3, u_{3\text{min}} - u_{3,t-1}, [0 0 1] \cdot P \cdot h_L) \]
\[ \beta_3 = \min(\Delta u_3, u_{3\text{max}} - u_{3,t-1}, [0 0 1] \cdot P \cdot h_H) \]

and
\[ P = \left( \begin{bmatrix} G'_1 & G'_2 & 0 \\ 0 & G'_3 & G'_4 \end{bmatrix} \right)^\top \left( \begin{bmatrix} G'_1 & G'_2 & 0 \\ 0 & G'_3 & G'_4 \end{bmatrix} \right)^{-1} \begin{bmatrix} G'_1 & G'_2 & 0 \\ 0 & G'_3 & G'_4 \end{bmatrix} \]

\[ h_L = \begin{bmatrix} y_{1\text{min}} - f_1 \\ y_{2\text{min}} - f_2 \end{bmatrix} \]
\[ h_H = \begin{bmatrix} y_{1\text{max}} - f_1 \\ y_{2\text{max}} - f_2 \end{bmatrix} \]

The constrained control problem in Equation 6.248 can be further expressed as an unconstrained control problem by including the Kuhn-Tucker multipliers as:
\[ \min_{\tilde{u}} J = \tilde{u}^\top H \tilde{u} + 2c^\top \tilde{u} + \mu^T A \]

where \( A \) is a vector of defined constraints:
\[ A = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6]^T \]

and \( \mu \) is a vector of Kuhn-Tucker multiplier associated with the constraints:
\[ \mu = [\mu_1 \ \mu_2 \ \mu_3 \ \mu_4 \ \mu_5 \ \mu_6]^T \]

Substituting Equations 6.250, 6.251, 6.257 and 6.258 into Equation 6.256 yields:
\[ \min_{\tilde{u}} J = h_{11} \Delta u_{1,t}^2 + (h_{21} + h_{12}) \Delta u_{1,t} \nabla u_{2,t} + h_{22} \nabla u_{2,t}^2 + (h_{32} + h_{23}) \nabla u_{2,t} \Delta u_{3,t} + h_{33} \Delta u_{3,t}^2 + 2c_1 \Delta u_{1,t} + 2c_2 \nabla u_{2,t} + 2c_3 \Delta u_{3,t} + \mu_1 a_1 + \mu_2 a_2 + \mu_3 a_3 + \mu_4 a_4 + \mu_5 a_5 + \mu_6 a_6 \]
The Kuhn-Tucker multipliers method requires that for $\tilde{u}$ to be optimal, it should satisfy the following conditions:

1. The first order conditions are:

$$\frac{\partial J}{\partial \tilde{u}} = 0$$  \hspace{1cm} (6.260)

2. The second order conditions are:

$$a_i \mu_i = 0 \text{ for all } i = 1 \text{ to } 6$$

such that

$$a_i = 0 \text{ if } \mu_i < 0$$  \hspace{1cm} (6.261)

or

$$a_i > 0 \text{ if } \mu_i = 0$$

The optimal solution $\tilde{u}$ of Equation 6.256 will be obtained when the 1st and 2nd order conditions are satisfied. The steps involved for the calculation of $\tilde{u}$ are given below:

**Step 1:** Determine an unconstrained control solution $\tilde{u}^*$.  

**Step 2:** Test to determine if any of the constraints in Equation 6.252 are violated. If all $a_i \geq 0$ (no constraints are violated) then go to **Step 5**.  

**Step 3:** Saturate the solution on one or two constraints and verify if the solution satisfies the above 1st and 2nd order conditions:

- Assume any one or two (say $a_m$ and $a_n$) of the constraints are satisfied by the constrained optimal solution, which gives a value of $\tilde{u}$. Therefore, all $\mu$'s other than $\mu_m$ and $\mu_n$ are zero.

- Calculate the Kuhn-Tucker multipliers $\mu_m$ and $\mu_n$ with the help of Equation 6.260. Verify if the Kuhn-Tucker multipliers $\mu_m, \mu_n \leq 0$
If the Kuhn-Tucker multipliers are negative, then go to Step 5.

**Step 4:** Repeat Step 3 for the next (set of) constraint(s).

**Step 5:** Implement the control action and repeat Steps 1-5 at next sample interval.

A total of 18 sets of constraints are possible in Step 3 of the above algorithm. They are listed in Appendix B.2.

**Optimal solution via the mixed-weights least-squares algorithm**

The optimal solution for the general case of constrained control problems can also be obtained through solving a mixed-weights least-squares problem. To do so, the quadratic cost function in Equation 6.245 is re-arranged as a 2-norm minimization problem as:

\[
\begin{align*}
\min_{\tilde{u}} J &= \left\| \begin{bmatrix}
    y_1 - w_1 \\
y_2 - w_2
\end{bmatrix} - \begin{bmatrix}
    \rho_1 I_{NU} & 0 & 0 \\
    0 & \rho_2 I_{NU} & 0 \\
    0 & 0 & \rho_3 I_{NU}
\end{bmatrix}^{1/2} \tilde{u} \right\|^2_2 \\
&= \left\| \begin{bmatrix}
    y_1 - w_1 \\
y_2 - w_2
\end{bmatrix} - \begin{bmatrix}
    \rho_1 I_{NU} & 0 & 0 \\
    0 & \rho_2 I_{NU} & 0 \\
    0 & 0 & \rho_3 I_{NU}
\end{bmatrix}^{1/2} \tilde{u} \right\|^2_2
\end{align*}
\]  

(6.262)

where \( \tilde{u} = [\tilde{u}_1^\top \tilde{u}_2^\top \tilde{u}_3]^\top \). Substituting Equation 6.240 into the above equation and modifying yields:

\[
\min_{\tilde{u}} J = \| R \tilde{u} - V \|^2_2
\]

(6.263)

where

\[
R = \begin{bmatrix}
    G_1' & G_2' & 0 \\
    0 & G_3' & G_4' \\
    \rho_1 I_{NU} & 0 & 0 \\
    0 & \rho_2 I_{NU} & 0 \\
    0 & 0 & \rho_3 I_{NU}
\end{bmatrix}^{1/2}
\]

(6.264)

and
The input amplitude constraints in Equation 6.247 can be modified as:

$$|L_u \tilde{u} - C_u| \leq 1$$  \hspace{1cm} (6.266)

or expressed as an $\infty$-norm problem as:

$$\|L_u \tilde{u} - C_u\|_\infty \leq 1$$  \hspace{1cm} (6.267)

where

$$L_u = \begin{bmatrix}
    A/u_{1,\text{radius}} \\
    0_{Nu} \\
    A/u_{3,\text{radius}}
\end{bmatrix} \quad C_u = \begin{bmatrix}
    (u_{1,\text{centre}} - u_{1,t-1})I_{NU \times 1}/u_{1,\text{radius}} \\
    0_{Nu \times 1} \\
    (u_{3,\text{centre}} - u_{3,t-1})I_{NU \times 1}/u_{3,\text{radius}}
\end{bmatrix} \hspace{1cm} (6.268)$$

and

$$u_{1,\text{centre}} = \frac{u_{1,\text{max}} + u_{1,\text{min}}}{2} \quad u_{1,\text{radius}} = \frac{u_{1,\text{max}} - u_{1,\text{min}}}{2}$$  \hspace{1cm} (6.269)

$$u_{3,\text{centre}} = \frac{u_{3,\text{max}} + u_{3,\text{min}}}{2} \quad u_{3,\text{radius}} = \frac{u_{3,\text{max}} - u_{3,\text{min}}}{2}$$  \hspace{1cm} (6.270)

Matrix $A$ is an $NU \times NU$ lower diagonal matrix whose non-zero entries are all equal to 1.

The output amplitude constraints in Equation 6.247 are modified in the form of:

$$|L_y \tilde{y} - C_y| \leq 1$$  \hspace{1cm} (6.271)

or expressed as an $\infty$-norm problem as:

$$\|L_y \tilde{y} - C_y\|_\infty \leq 1$$  \hspace{1cm} (6.272)
where

\[ L_y = \begin{bmatrix} G_1'/y_1,\text{radius} & G_2'/y_1,\text{radius} & 0_{N_y \times N_u} \\ 0_{N_y \times N_u} & G_3'/y_2,\text{radius} & G_4'/y_2,\text{radius} \end{bmatrix} \]

(6.273)

\[ C_y = \begin{bmatrix} (y_{1,\text{centre}} I_{N_y \times 1} - f_1)/y_1,\text{radius} \\ (y_{2,\text{centre}} I_{N_y \times 1} - f_2)/y_2,\text{radius} \end{bmatrix} \]

and

\[ y_{1,\text{centre}} = \frac{y_{1,\text{max}} + y_{1,\text{min}}}{2} \quad y_{1,\text{radius}} = \frac{y_{1,\text{max}} - y_{1,\text{min}}}{2} \]

(6.274)

\[ y_{2,\text{centre}} = \frac{y_{2,\text{max}} + y_{2,\text{min}}}{2} \quad y_{2,\text{radius}} = \frac{y_{2,\text{max}} - y_{2,\text{min}}}{2} \]

(6.275)

The input rate constraints in Equation 6.247 are re-arranged in the form:

\[ |L_{du} \tilde{u} - C_{du}| \leq 1 \]

(6.276)

or expressed as an \( \infty \)-norm problem as:

\[ ||L_{du} \tilde{u} - C_{du}||_{\infty} \leq 1 \]

(6.277)

where

\[ L_{du} = \begin{bmatrix} I_{N_u}/\Delta u_1 & 0_{N_u} & 0_{N_u} \\ 0_{N_u} & A_1/\Delta u_2 & 0_{N_u} \\ 0_{N_u} & 0_{N_u} & I_{N_u}/\Delta u_3 \end{bmatrix} \quad C_{du} = \begin{bmatrix} 0_{N_u \times 1} \\ A_2/\Delta u_2 \\ 0_{N_u \times 1} \end{bmatrix} \]

(6.278)

and

\[ A_1 = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} (\Delta u_{2,t-1} - \Delta u_{2,t})/\Delta u_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]

(6.279)
Combining Equations 6.267, 6.272 and 6.277 gives:

\[ \| \mathbf{L} \hat{\mathbf{u}} - \mathbf{C} \|_\infty \leq 1 \]  

(6.280)

where

\[
\mathbf{L} = \begin{bmatrix} L_u \\ L_y \\ L_{du} \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} C_u \\ C_y \\ C_{du} \end{bmatrix}
\]

(6.281)

The constrained control problem is then solved by minimizing the cost function in Equation 6.263 with the satisfaction of the constraints in Equation 6.280, i.e.

\[
\min_{\hat{\mathbf{u}}} J = \| \mathbf{R} \hat{\mathbf{u}} - \mathbf{V} \|_2^2
\]

subject to:

\[ \| \mathbf{L} \hat{\mathbf{u}} - \mathbf{C} \|_\infty \leq 1 \]  

(6.282)

The optimal solution of the above constrained problem can be obtained through minimizing the following mixed-weights least-squares or mixed-objective function:

\[
\min_{\hat{\mathbf{u}}} \left\| \left( \begin{bmatrix} \mathbf{W}^{(i)}_{f} \\
\end{bmatrix} \right)^{1/2} \left( \begin{bmatrix} \mathbf{R} \\
\mathbf{L} \\
\end{bmatrix} \hat{\mathbf{u}} - \begin{bmatrix} \mathbf{V} \\
\mathbf{C} \end{bmatrix} \right) \right\|_2^2
\]

(6.283)

where \( w^{(i)} \) is a positive real scalar and \( W^{(i)} \) is a diagonal matrix with positive entries. Both \( w^{(i)} \) and \( W^{(i)} \) are calculated iteratively as:

\[
w^{(i+1)} = \frac{w^{(i)}}{\sum_{k=1}^{m} w_{kk}^{(i)} |e_k^{(i)}|} \quad W^{(i+1)}_{jj} = \frac{W_{jj}^{(i)} |e_j^{(i)}|}{\sum_{k=1}^{m} w_{kk}^{(i)} |e_k^{(i)}|}
\]

(6.284)

where

\[ e^{(i)} = \mathbf{L} \hat{\mathbf{u}}^{(i)} - \mathbf{C} \]

(6.285)

and

\[ e^{(i)} = [e_1^{(i)} e_2^{(i)} \ldots e_m^{(i)}]^T \]

(6.286)

\( W_{jj}^{(i)} \) are the entries of the \( W^{(i)} \) and \( m \) is the number of the constraints (\( m \leq 2NU + N_2 \)).
6.3.3 Simulation Results and Discussions

The proposed control strategy described previously was tested on a combination of empirical and mechanistic models (see Figure 6.43), to approximate the features of an industrial situation. The dynamics from inputs to outputs was based on the identification results performed on industrial refiners. The static relationship between inputs and outputs was simulated based on the mechanistic models. The effect of operating variables on refining disturbances

\[ \begin{align*}
\text{Pc} & \quad \text{1st order + delay} \\
\text{St} & \quad \text{1st order + delay} \\
\text{Fd} & \quad \text{1st order + delay}
\end{align*} \]

the primary refiner

E : specific energy 
Fc: production rate 
P : motor load 
c : outlet consistency 
\( \tau \) : residence time 
I : refining intensity 
N : number of bar impacts 
PQ: pulp quality

Figure 6.43: Input-output dynamic and static relationship for the primary refiner

consistency was modelled using mass and energy balances. Production rate control was operated at steady state and it was reasonably assumed to be a linear function of the transfer screw speed.

The model used for the controller design was given in Equation 6.240. In the following simulations, a nominal set of tuning parameters \( q_1 = q_2 = 1, N_1 = 1, N_2 = 15, NU = 1 \) and \( \lambda_1 = \lambda_2 = 0.98 \) were used.

Response to Rate Constraints
Due to equipment physical limits, manipulated variables should be bounded before their implementation. Figure 6.44 shows constrained control via MWLS with the rate constraints on manipulated variables. The hard limits on the closing pressure, screw speed and dilution flow were set to $|\Delta u_1| \leq 50, |\Delta u_2| \leq 0.2$ and $|\Delta u_3| \leq 0.5$ respectively. The controller parameters were $\rho_2 = 1.0, \rho_1 = \rho_3 = 0.0001$. No constraints on the amplitudes of the manipulated variables were implemented. To compare constrained control with unconstrained one, the simulation was run without hard constraints for the first 200 intervals and constrained control was turned on at interval 200. Figure 6.44 shows the movements of the manipulated variables were effectively reduced into their desired limits.

**Response to Amplitude Constraints**

Amplitudes of manipulated variables should also be limited to their desirable ranges for the reasons of safety, stable operation and equipment physical limitation. This is particularly important in a TMP process, e.g. maximum closing pressure applied for each refiner to avoid plate clashes, maximum and minimum dilution flow for the reasons of stable operation and dilution valve’s physical limit. Figure 6.45 shows the simulation performed without constraints for the first 200 sec. and with hard constraints for the rest of time. The amplitude constraints on the closing pressure $u_1$ and dilution flow $u_3$ were set to $820 \leq u_1 \leq 1150$ and $3.4 \leq u_3 \leq 5.2$. Figure 6.45 demonstrates that the setpoints were achieved by manipulating closing pressure and dilution flow within their desirable ranges while more room for manipulation on the screw speed to compromise the outputs deviations from the targets.

**Response to Deterministic Disturbances and Changes in Process Dynamics**

Figure 6.46 shows the response to a number of deterministic load disturbances and changes in the process dynamics. Six time series plots are shown in the figure: (i) the motor load ($y_1$) and its setpoint change, (ii) the consistency ($y_2$) and its setpoint, (iii)
Figure 6.44: Constrained control via MWLS, $t \geq 200$ sec: $|\Delta u_1| \leq 50$, $|\Delta u_2| \leq 0.2$, $|\Delta u_3| \leq 0.5$ ($q_1 = q_2 = 1$, $p_2 = 1.0$, $p_1 = p_3 = 0.0001$)
Figure 6.45: Constrained control via analytical QP, $t \geq 200$ sec: $820 \leq u_1 \leq 1150$, $3.4 \leq u_3 \leq 5.2$ ($q_1 = q_2 = 1$, $p_2 = 1.0$, $p_1 = p_3 = 0.001$)
the closing pressure \((u1)\), (iv) the target and actual screw speed \((u2)\), (vi) the dilution flow \((u3)\), and (vii) the target and actual production rate. A set of tuning parameters \(\rho_2 = 10, \rho_1 = \rho_3 = 0.0001\) were used with the objective of maintaining the production rate to its target while manipulating the closing pressure and dilution to reach the setpoints.

Load disturbances both from the motor load to the consistency and the dilution to the motor load were simulated in Figure 6.46. An increase (or decrease) in the motor load results in an increase (or decrease) in the consistency. This is because increased (or decreased) motor load results in more (or less) water converted into steam and therefore, an increase (or decrease) in the consistency. To compensate for the motor load effect on the consistency, the dilution flow was adjusted to maintain the target of the consistency. Due to process nonlinearity, an adaptive scheme was employed to estimate time-varying gain on-line. Figure 6.46 shows that proposed controller performed well with process nonlinearity and time-varying gain.

**Tuning to Reduce Chip Screw Manipulation**

Figure 6.47 shows a simulation performed with the objective of decreasing the amount of chip screw manipulation while maintaining the targets of the motor load and consistency. This may be considered a realistic control objective because excessive screw speed manipulation can result in a large variation in production rate and thereby, varying the specific energy applied and pulp quality. To reduce screw manipulation, \(\rho_2\) was set to \(\rho_2 = 10\). Figure 6.47 shows substantial reduction in the screw speed by increasing associated weighting factor, without sacrificing output performance.

**The pulp quality response to setpoint changes**

Figures 6.48 and 6.49 show a set of simulation runs undertaken with the changes in the motor load setpoint. The initial motor load setpoint was determined from given total
Figure 6.46: Unconstrained controller \( (q_1 = q_2 = 1, \rho_2 = 1.0, \rho_1 = \rho_3 = 0.0001) \)
Figure 6.47: Unconstrained controller \((q_1 = q_2 = 1, \rho_2 = 10, \rho_1 = \rho_3 = 0.0001)\)
specific energy, power split and production rate as follows:

\[
\begin{align*}
\text{Total specific energy} & = 6.0(MJ/kg) \\
\text{Power split} & = 57/43 \\
\text{Production rate} & = 325(t/d)
\end{align*}
\]

The consistency setpoint was set to 45%. These setpoints are given according to the requirement of product quality. Perturbation signals were added at each control output to obtain persistent excitation due to an adaptive scheme applied. The tuning parameters were set to \( q_1 = q_2 = 1, \rho_2 = 1, \rho_1 = \rho_3 = 0.0001 \). An increase in the motor load setpoint was made at \( t = 150\text{sec.} \) and a decrease in the consistency setpoint was made at \( t = 300\text{sec.} \). Figure 6.48 shows that the motor load setpoint was achieved through manipulating the closing pressure while the screw speed was adjusted around desired production rate. In practice, an increase in the motor load will increase the consistency. Figure 6.48 shows that consistency variations resulting from the load setpoint change was compromised through automatically increasing the dilution flow.

Figure 6.49 shows that at the given production rate, an increase in the motor load increased the specific energy. As a result, the inlet consistency decreased for a given outlet consistency because with higher specific energy, there will be more water converted to steam and therefore, less water in the refining zoom. To keep the consistency in the refining zone constant, inlet dilution was increased and as a result the inlet consistency reduced. The residency time is proportional to the specific energy and the inlet consistency. Increased residence time indicated that an increase in the specific energy was relatively significant than a decrease in the inlet consistency. Increases in the refining intensity and the specific power were due to a relatively significant increase in the specific energy. A significant decrease in the freeness was resulted from a decrease in the specific energy. Increased refining intensity and specific power due to an increase in the specific energy resulted in decreases in the long fibre and shive contents.
Figure 6.48 shows that a decrease in the consistency setpoint was made at $t = 300\ sec$. With the proposed control, the consistency setpoint was achieved by increasing the dilution flow and slightly decreasing the chip feed due to a tight constraint on the production rate. Figure 6.49 shows that for a given specific energy, decreased consistency resulted in decreases in the inlet consistency and the residence time whereas increases in the refining intensity and the specific power. The explanation is as follows. Increased dilution flow at a given specific energy brought about a higher centrifugal force on the pulp flow. As a result, the residence time for a pulp to pass through the refining zone was reduced and thereby, the total number of bar impacts received by the pulp was reduced. At a given specific energy, reduced number of bar impacts increased the average specific energy per impact, i.e., the refining intensity. Increased refining intensity and specific power caused more fibre cutting, resulting in less long fibre content and shive content. Due to a strong relationship with the specific energy, the freeness was almost unchanged due to a constant specific energy.
Figure 6.48: Unconstrained controller ($\rho_2 = 1, \rho_1 = \rho_3 = 0.0001$)
Figure 6.49: Unconstrained controller ($\rho_2 = 1, \rho_1 = \rho_3 = 0.0001$)
6.4 Laguerre Function-Based Control of TMP Refiners

Control of a wood chip refiner in TMP manufacturing is sometimes difficult due to the incremental gain between refiner motor load and plate gap subject to a sudden change in the sign. This is particularly critical when the refiner is operating close to the maximum load point (see Figure 6.50). The linear approximation-based control for this case may not be satisfactory and may lead to unstable operation. Alternatively, the Laguerre functions presentation is used in the following to approximate process dynamics and static nonlinearity. A generalized predictive control algorithm is then derived based on the nonlinear Laguerre function representation.

6.4.1 Nonlinearity Description

It has been known for a long time that the incremental gain (see Figure 6.50) between refiner motor load and plate gap is subject to a slow drift due to plate wear and to
sudden changes in sign due to collapse of the pulp pad in the refining zone. Plate wear generally occurs gradually over a period of several hundred hours, but can also occur suddenly in the case of plate clash, i.e., metal to metal contact between two plates. The maximum load corresponds to the minimum gap below which the pulp pad in the refining zone cannot sustain the pressure and then collapses, resulting in plate clashes and the incremental gain between the load and the gap changing in sign from negative to positive. A plate clash is highly undesirable because it will disrupt production and damage the plates. To avoid a plate clash, the gap has to be opened beyond the point where the collapse occurred, i.e., the refiner has to be operating in the region where closing the gap increases the load. However, the gap at which the pad collapses is unpredictable, and it depends on wood species, wood chip quality, refining consistency, plate condition, etc. These make refiner control more difficult.

6.4.2 Past Control Solutions

A first attempt at applying adaptive control to the problem was made by Dumont [30]. In his solution, Fortescue's variable forgetting factor based recursive least squares estimator was used to identify the gain, and a set of rules were set to back out the refiner plates and restore the pulp pad when the gain changed sign. Some success on an industrial chip refiner was reported, but the choice of the variable forgetting factor design parameters to achieve the required trade-off between tracking slow and fast gain changes proved to be difficult in practice [33]. Dumont and Åström [32] then investigated several alternatives including dual control and proposed use of a nonlinear performance index which explicitly penalizes operation in the pad collapse region. In their proposed control strategy, additional terms reflecting the nonlinear nature of the process were introduced into a performance index, instead of using an index function reflecting the output error alone. Using this control criterion, penalty is given when the estimated gain is small. In
the case of an unreachable motor load setpoint, the controller will tend to keep the gain negative instead of eliminating the output error. When the motor load is lower than the setpoint, such control will lead the control action to close the gap when the estimated gain is negative with relatively large amplitude and open the gap when gain tends to be positive. Their idea was further developed and tested on an industrial refiner by Allison et al. [48]. Preliminary industrial results showed the general success and the potential of industrial control implementation. However, one drawback is the computational difficulty attached to solving the dual control problem on-line. A new way of representing a refiner was proposed by Dumont and Fu [114] for the control of a wood chip refiner. In their control scheme, the Laguerre function expansion via Volterra kernel was used to model process nonlinearity and dynamics, and the clashes of the refiner plate was avoided without adding extra measurements and computational difficulty which are associated with dual control problem.

The control approaches given above were based on single-variable control philosophy (i.e., the pulp quality is determined by the energy input), so they were limited by their control performance. This is because pulp quality is mainly a function of two variables: the specific energy and the specific energy per bar impact or refining intensity. For a given plate configuration and a constant disc rotational speed, the refining intensity is, with some approximation, largely dependent on refining consistency. To obtain desired pulp quality, the specific energy as well as the refining consistency need to be controlled. The following discussion presents a multivariable Laguerre function-based control scheme in which the specific energy and the refining intensity are controlled through control of the motor load and refining consistency by manipulating the hydraulic closing pressure to position the gap and the dilution flow rate.
6.4.3 2 × 2 Mixed Linear-Nonlinear Laguerre Function-Based Control Of TMP Refiners

A 2 × 2 Laguerre function-based predictive control scheme developed in Section 5.4 is applied towards stable control of a TMP refiner when the refiner is operating close to the maximum load.

![Diagram of 2x2 Laguerre model for TMP refiner]

Figure 6.51: Structure of a 2×2 Laguerre model for a TMP refiner

**Laguerre Model**

Figure 6.51 shows the 2 × 2 Laguerre model structure using the first $N$ filters for modelling a wood chip refiner. The nonlinear dynamic relationship between the motor load and plate gap is modelled using the nonlinear Laguerre functions representation, whereas the dynamic relationship between the outlet consistency and dilution flow rate is approximated by the linear Laguerre functions representation. The mathematical formula of this Laguerre model structure was given in Equation 5.205.

**Adaptive-Predictive Control based on Laguerre Model**
A generalized predictive control algorithm is then derived based on the mixed linear-nonlinear Laguerre model. The control objective is to maintain the motor load $y_1$ and refining consistency $y_2$ to their perspective setpoints according to the desired specific energy and refining intensity by adjusting the plate gap through manipulating the hydraulic closing pressure $u_1$, and the dilution flow rate $u_2$ to the refiner. The control actions are calculated by minimizing the performance index given in Equation 5.213. The Laguerre model given in Equations 5.205 is used to predict process output trajectories. The differences between the predictions and actual process measurements are used to estimate the model parameters on-line to handle the time-varying nature of the process.

6.4.4 Simulation Results and Discussion

The described adaptive-predictive control scheme based on the mixed linear-nonlinear Laguerre model has been tested extensively in simulations. The simulation results show its potential applicability and illustrate its practicality as follows. The simulations were performed on a combination of empirical and mechanistic models given in Equations 6.287 ~ 6.290, to approximate the features of an industrial situation. The dynamics from inputs to outputs were based on the identification results performed on industrial refiners. The static relationship between the plate gap and the motor load was simulated as shown in Figure 6.50. The effect of operating variables on the consistency was simulated using the results based on the mass and energy balances. The above described simulation models may be presented as:

$$ x_{1,t} = a_{11} x_{1,t-1} + b_{11} u_{1,t-2} + b_{12} u_{1,t-3} \tag{6.287} $$

$$ x_{2,t} = a_{21} x_{2,t-1} + b_{21} u_{2,t-2} + b_{22} u_{2,t-3} \tag{6.288} $$

$$ y_{1,t} = g_1 - g_2 \frac{(\alpha x_{1,t})^4 + 1/3}{(\alpha x_{1,t})^3} + N_{1,t} \tag{6.289} $$
Chapter 6. Constrained Multivariable Control Of A TMP Plant

\[ y_{2,t} = \frac{F_c L}{E_s L + x_{2,t} L - \eta y_{1,t}} + N_{2,t} \]  

(6.290)

where \( y_1 \) is the motor load, \( y_2 \) is the outlet consistency, \( u_1 \) is the pulse sent to the hydraulic cylinder to position the gap \( x_1 \), and \( u_2 \) is the dilution flow rate used to control the consistency. \( N_i \) in the above equation account for the joint effect of all external disturbances and measurement noise on an output. In the following simulations, the linear dynamic parameters in the above equations are: \( a_{11} = 0.75, b_{11} = 0.25, b_{12} = 0, a_{21} = 0.8187, b_{21} = 0.1813, b_{22} = 0 \). The nonlinear relationship given in Equation 6.289 is simulated with \( g_1 = 7600, g_2 = 800, \alpha = 1 \) in a normal state, and \( \alpha = 0.77 \) in a collapsed state when \( x < 0.9 \). Other operating variables are: the pulp production rate \( F_c = 3.7616 \text{ kg/sec} \), the latency heat \( L = 2.5 \text{ MJ/kg} \), the inlet solid content \( s = 0.5 \) and the power efficiency coefficient \( \eta = 0.9 \). However, those parameters and nonlinear functions in the above equations are supposed unknown. There are white noises added to the motor load and consistency outputs, simulating measurement noises.

The predictive controller tuning parameters, in the following simulations, are set to \( N_1 = 1, N_2 = 15, \rho_1 = 0.0 \) and \( \rho_2 = 0.0001 \). The amplitude and rate constraints on the plate gap are applied, which are \( 0.7 \leq x_{1,t} \leq 5.0 \) and \( |\Delta x_{1,t}| \leq 0.1 \). To start with simple Laguerre functions representation, two Laguerre filters are used with the Laguerre filter poles \( a_1 = a_2 = 0.7 \). The severe nonlinear relationship between the plate gap and the motor load is modelled by the 2nd order nonlinear Laguerre function. The RLS estimation algorithm is used with the forgetting factors: \( \lambda_1 = \lambda_2 = 0.98 \).

**Motor Load Setpoint Unreachable.** Figures 6.52 shows a simulation of the proposed controller when the load setpoint is unreachable. For the period up to about \( t = 80 \) sec, the refiner is operating in the normal state, i.e., closing the gap increases the motor load. The controller for this case is able to adjust the plate gap through manipulating the closing pressure to minimize the error between the load and its setpoint. The load
setpoint then begins to increase at \( t = 80 \) sec from 6380 kW to 6580 kW. The controller, in the beginning, closes the plate gap in order to reach the load setpoint. As the gap decreases, the extreme point (see Figure 6.50) will be reached, because of the setpoint is unreachable in this case. The controller based on the estimated nonlinear Laguerre model is able to identify the extreme point and evaluate the control input close to the true extreme point 0.9. Correspondingly, the motor load output is then stabilized below its setpoint.

*Plate Clash Avoided.* Figure 6.53 demonstrates how a plate clash can be avoided by the proposed control strategy. As can be seen at about \( t = 350 \) sec in Figure 6.53, when a refiner is operating near the maximum load point, just process disturbances can produce a pulp pad collapse. As a result, the refiner will be operating on the left side of the extreme point (see Figure 6.50), i.e., closing the gap decreases the motor load. For this case, the nonlinear Laguerre model based controller is able to detect this collapse and quickly open the gap, allowing the pulp pad to rebuild. The refiner then goes back to the normal operation state and a plate clash is avoided.

*Changes in Consistency Setpoint.* Figure 6.54 shows a simulation of the proposed control scheme when the consistency setpoint changes. As can be seen, a step change in the consistency setpoint is made at \( t = 220 \) sec and the proposed controller is able to adjust the dilution flow rate to meet the setpoint. Although the effect of operating variables on the consistency is nonlinear (Equation 6.290), the proposed controller based on the estimated linear Laguerre approximation is able to identify the nonlinearity as a gain change and evaluates the control input accordingly.

*Response to Load Disturbances.* In practice, there is a coupling between the motor load and refining consistency, i.e., an increase in the motor load will increase the consistency and vice versa. Figure 6.52 shows that to maintain the outlet consistency constant when the motor load increases resulting from an increase in its setpoint, the dilution flow rate is
increased. A decrease in the motor load will decrease the outlet consistency. To maintain a constant consistency, the proposed controller automatically reduces the dilution flow to the refiner. Figure 6.54 shows an increase in the outlet consistency will result in an increase in the load. To keep the load constant, some control effort goes into decreasing the load which is done by opening the plate gap.

Simulations of the proposed control scheme demonstrated that it is more robust than the conventional model-based approaches in the sense of no need for assumption of the model order and time-delay, and is able to handle an industrial process with severe nonlinearity, unknown dynamics and external disturbances. In the case of a pulp pad collapse, the proposed controller was able to identify the collapse and quickly open the gap for the pad to rebuild, whereas the linear model-based control schemes for this case may cause some control problems such as process oscillation, pulp pad collapse acceleration and plate clashes. Using Laguerre functions approximation, plant nonlinearity and dynamics can be modelled properly with a flexible model structure and less model parameters.

6.5 Summary and Conclusions

A new control strategy was developed towards better control of a two-stage TMP plant, where pulp quality was controlled through selecting the setpoints of the specific energy and the refining intensity at each stage. The targets of the specific energy and the refining intensity were achieved through control of the motor load, production rate and outlet consistency by manipulating the closing pressure, chip feed and dilution flow rate. The proposed control strategy is advantageous because it defines chip refining more accurately than conventional quality-energy control schemes, i.e., chip refining was defined by the specific energy alone. Therefore, the proposed control scheme provides more degrees of freedom in the control of the process. A constrained adaptive-predictive controller based
Figure 6.52: Motor load setpoint unreachable
Figure 6.53: A plate clash was avoided by the proposed control scheme
Figure 6.54: Response to the changes in consistency setpoint
on non-square MIMO CARIMA model was developed for the control of TMP refiners in the process. The control algorithm was based on the GPC because of its simplicity and ease of use along with its ability to handle problems in one algorithm. An optimal solution of the constrained control problem was obtained by solving a QP problem. For the case of unfeasible constrained optimization problem, the MWLS was used as an alternative to calculate an optimal solution of the constrained control problem. Simulations of the proposed control schemes showed the successes in handling problems such as coupling, variable interactions, time-varying dynamics, time delay and external disturbances in one algorithm, without violations of input and output constraints.

Control of a wood chip refiner in a TMP process is sometimes difficult due to the severe nonlinearity of the process, i.e., the incremental gain between refiner motor load and plate gap is subject to a sudden change in the sign. This is particularly critical when the refiner is operating close to the maximum load point. The linear approximation-based control for this case may lead to unstable operation. To avoid the problem, the nonlinear Laguerre functions were used to approximate process static nonlinearity and dynamics. The multivariable generalized predictive controller is then derived based on a mixed linear-nonlinear Laguerre model. The nonlinear Laguerre functions were used to represent the severe nonlinearity between refiner motor load and plate gap, while the linear Laguerre functions were used to approximate nonlinearities that are not severe. For an adaptive control implementation, the nonlinear Laguerre model can be arranged in a linear form in the parameters so that the powerful but simple Recursive Least Squares (RLS) algorithm can be used for the parameter estimation. Simulations of the proposed control scheme showed its successes in setpoint tracking as well as load disturbance reject. In the case that the incremental gain between refiner motor load and plate gain was subject to a sudden change in the sign due to a pulp pad collapse, the proposed controller had the ability to detect the collapse and quickly open the gap for the pad to
rebuild in order to avoid a plate clash.
Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis has presented the development of a constrained multivariable predictive control approach to the control of a TMP plant. Control of this particular process is difficult because of the multivariable nature, process constraints, nonlinearity, time-varying dynamics and time delay along with stochastic disturbances. The objective of this work, therefore, has been to develop a control strategy suitable for handling these problems and for an industrial implementation towards an automated control of TMP manufacturing.

The process was modelled using a combination of mechanistic and empirical methods. Static relationships between inputs and outputs were obtained based on first principles such as mass and energy balances to represent the quantitative relationships and complex interactions between variables. To avoid the complexity of the development of differential equations, the dynamics of the process were obtained through process identification. Mill experiments were performed in Canadian mills from which a large number of on-line operational data was collected. Identification results showed that process dynamics can be reasonably represented by first order with time delay and that process disturbances feature nonstationary characteristics with first-order dynamics. The developed model was used for process pre-analysis, controller design and proposed controller investigation.

Instead of the usual quality-energy control approaches, pulp quality was considered as a function of two variables: the energy per unit mass of pulp or specific energy and the
specific energy per bar impact or refining intensity. Therefore, pulp quality was controlled through selecting the setpoints of specific energy and refining intensity for each stage. The targets of specific energy and refining intensity at each stage were achieved through the control of motor load, production rate and refining consistency. The manipulated variables for obtaining the desired operating conditions were chip feed, dilution flow rate and closing pressure at the inlet of each refiner.

In TMP manufacturing, the constraints on different flows, as well as constraints involved by equipment physical limitations and product quality requirements, all play a crucial role in control design. Applications of control action without consideration of the constraints may lead to a degraded process performance and, in some cases, closed-loop instability. To deal with the problems created by constraints, a constrained multivariable adaptive-predictive control scheme was developed for the control of a wood chip refiner of the TMP process. The control algorithm was based on GPC because it is easily understood and can handle the problems in one algorithm along with its ability to incorporate constraints explicitly. Future control actions were determined at each sampling time by minimizing the predicted errors subject to operating constraints. On-line estimates of a MIMO CARIMA model were used for output prediction. The model structure was based on the identification results obtained from the process identification stage. An analytical solution of the constrained MIMO GPC for $NU = 1$ was derived by solving a quadratic programming problem. In the case of unfeasible constraints, the mixed-weights least-squares algorithm was used to solve the constrained optimization problem. The MWLS-based solution will always converge to the point that minimizes the maximum constraint violation. The proposed control schemes were investigated on a combination of empirical and mechanistic models obtained from the modelling stage to approximate an industrial situation. Simulation results showed that the proposed schemes were able to handle the problems in one algorithm without violation of process constraints along
Chapter 7. Conclusions and Future Work

with the potential of optimizing process operation and improving product quality.

Control of a wood chip refiner is challenging due to the problem associated with the severe nonlinearity between refiner motor load and plate gap, i.e., the incremental gain between refiner motor load and plate gap is subject to a sudden change in the sign due to pulp pad collapse. This problem is particularly critical when the refiner is operating close to the maximum load point. To deal with the problem, a new approach was taken to represent the process, i.e., a MIMO Laguerre model was used to approximate process dynamics as well as static nonlinearity. Generalized predictive control solutions were then derived based on a mixed linear-nonlinear Laguerre function representation. The nonlinear Laguerre functions were used to approximate the severe nonlinearity between the refiner motor load and plate gap, whereas the linear Laguerre functions were used to represent nonlinearities that are not severe. The Laguerre function-based control scheme is advantageous because of its simplicity and flexibility while eliminating the need for assumptions about model order and time delay. Simulations of the proposed control scheme demonstrated that the nonlinearity and dynamics of the process can be represented properly even using lower order Laguerre functions. When the sign of the incremental gain between load and gap started to change, the proposed controller was able to detected the change and quickly open the gap to avoid plate clashes. The Laguerre model-based control scheme was considered as an alternative to the transfer function-based control when a wood chip refiner is operating close to the maximum load point.

7.2 Future Work

It was difficult to address all the issues that arose during this research. A number of topics still need to be dealt with in the future.
The proposed control schemes have been tested on a combination of empirical and mechanistic models to approximate an industrial plant. Further research could be undertaken towards in-depth investigation of the behaviour of the proposed control schemes over two stages in series.

Throughout the study and simulation in this thesis, desired pulp quality was achieved through control of specific energy and refining intensity by adjusting the setpoints of motor load, throughput and outlet consistency. Further work would be undertaken by including quality feedback control. An outer loop based on measurements of pulp properties such as freeness, long fibre content and shive content would be included as a cascade control in providing the setpoints for inner loop control of specific energy and refining intensity at each stage.

In this thesis, the chip screw speed was constrained about the desired production rate while still leaving some room for manipulation. An extension of this study would consider maximizing the production rate for a given energy input to meet the different requirements of the process.
Bibliography


Appendix A

Time Series Analysis Tools

A.1 Auto-Correlation Functions

A stationary Gaussian time series \( y_t \) can be uniquely described by its mean \( \bar{y} \) and auto-covariance \( c_{yy}(k) \), or equivalently by its mean \( \bar{y} \), variance \( \sigma_y^2 \) and auto-correlation function \( r_{yy}(k) \). The auto-correlation function \( r_{yy}(k) \) of \( y_t \) at lag \( k \) is defined as:

\[
r_{yy}(k) = \frac{c_{yy}(k)}{\sigma_y^2}
\]

where

\[
c_{yy}(k) = \frac{1}{N} \sum_{t=1}^{N-k} (y_t - \bar{y})(y_{t+k} - \bar{y})
\]

\[
\sigma_y^2 = \frac{1}{N} \sum_{t=1}^{N} (y_t - \bar{y})^2
\]

\[
\bar{y} = \frac{1}{N} \sum_{t=1}^{N} y_t
\]

where \( c_{yy}(k) \) is auto-covariance of \( y_t \) at lag \( k \). \( k = 0, 1, 2, \cdots, K \) is the number of discrete sampling intervals between the observations of \( y_t(t = 1, 2, \cdots, N) \). \( \sigma_y^2 \) is the special case of \( c_{yy}(k) \) at \( k = 0 \) (i.e., \( \sigma_y^2 = c_{yy}(0) \)) and it presents the variance of the time series \( y_t \). It can be seen from above equation that \( c_{yy}(k) = c_{yy}(-k) \) and \( r_{yy}(0) = 1 \) (i.e., auto-correlation is always unity at \( k = 0 \)). Therefore, the auto-correlation \( r_{yy}(k) \) is often plotted for the value of the lag \( k = 0, 1, \cdots, K \).

The auto-correlation coefficients \( r_{yy}(k) \) generally tend to zero as \( k \) increases. To have a
Appendix A. Time Series Analysis Tools

A crude check on whether auto-correlation is effectively zero beyond a certain lag, the auto-correlations are usually plotted out with their 95% confidence intervals. An approximate expression for the variance of estimated $r_{yy}(k)$ is defined in the following [75] on the assumption that the time series $y_t$ is completely random (uncorrelated or white).

$$\text{var}[r_{yy}(k)] \simeq \frac{1}{N}$$  \hspace{1cm} (A.293)

Thus standard error (S.E.) of the estimated partial auto-correlation is

$$S.E.[r_{yy}(k)] \simeq \frac{1}{\sqrt{N}}$$  \hspace{1cm} (A.294)

Thus 95% confidence intervals are given by

$$95\% \text{ Confidence Intervals} \pm 2S.E. = \frac{2}{\sqrt{N}}$$  \hspace{1cm} (A.295)

Any auto-correlation coefficient $r_{yy}(k)$ which falls in 95% confidence intervals is insignificant.

A.2 Partial Auto-Correlation Functions

Partial auto-correlation function is a tool used to identify the order of an Auto-Regressive (AR) function which is fitted by an observed time series. The partial auto-correlation function $\phi_{yy}^{kk}$ of a time series $y_t$ at lag $k$ may be estimated by fitting successively AR processes of order $1, 2, \ldots, k$ shown in the following:

$$r_{yy}(j) = \phi_{yy}^{k1} r_{yy}(j-1) + \cdots + \phi_{yy}^{k(k-1)} r_{yy}(j-k+1) + \phi_{yy}^{kk} r_{yy}(j-k)$$  \hspace{1cm} (A.296)

$$j = 1, 2, \ldots, k$$
Appendix A. Time Series Analysis Tools

Solving above equations for \( k = 1, 2, 3, \ldots \) successively by applying a simple recursive method and picking out the estimates \( \phi_{yy}^{11}, \phi_{yy}^{22}, \ldots, \phi_{yy}^{kk} \). For an AR process of order \( n \), the partial auto-correlation coefficient \( \phi_{yy}^{kk} \) will be:

\[
\begin{align*}
\phi_{yy}^{kk} & \neq 0 \quad \text{if} \quad k \leq n \\
\phi_{yy}^{kk} & = 0 \quad \text{if} \quad k > n
\end{align*}
\]  
(A.297)

In other words, the partial auto-correlation function of a \( n \)-th-order AR process has a cutoff after lag \( k = n \).

The partial auto-correlation coefficients \( \phi_{yy}(k) \) generally tend to zero as \( k \) increases. To have a crude check on whether the partial auto-correlation function is effectively zero beyond a certain lag, the partial auto-correlation coefficients are usually plotted out with their approximate 95% confidence intervals. The approximate 95% confidence intervals on the partial auto-correlations are defined as the same as in Equation A.295 under the assumption that the series \( y_t \) is white.

A.3 Cross-Correlation Functions

The cross-correlation function \( r_{uy}(k) \) between input \( u_t \) and output \( y_t \) is a tool used to model the input-output dynamics of the process by fitting the time series pair \( (u_t, y_t) \) to a model. The cross-correlation function \( r_{uy}(k) \) is estimated by

\[
r_{uy}(k) = \frac{c_{uy}(k)}{\sigma_u \sigma_y}
\]  
(A.298)

where

\[
c_{uy}(k) = \begin{cases} 
\frac{1}{N} \sum_{t=1}^{N-k} (u_t - \bar{u})(y_{t-k} - \bar{y}), & k = 0, 1, 2, \ldots \\
\frac{1}{N} \sum_{t=1}^{N+k} (y_t - \bar{y})(u_{t-k} - \bar{u}), & k = 0, -1, -2, \ldots 
\end{cases}
\]  
(A.299)
Appendix A. Time Series Analysis Tools

$c_{uy}(k)$ is the cross-covariance coefficients between $u_t$ and $y_t$ ($t = 1, 2, \ldots, N$) at lag $k$. $\bar{u}$ and $\bar{y}$ are the means of the series $u_t$ and $y_t$ respectively. $\sigma_u^2$ and $\sigma_y^2$ are the variances of the series $u_t$ and $y_t$ respectively. $\sigma_u^2$ is the special case of $c_{uu}(k)$ at $k = 0$ (i.e., $\sigma_u^2 = c_{uu}(0)$), while $\sigma_y^2$ is the special case of $c_{yy}(k)$ at $k = 0$ (i.e., $\sigma_y^2 = c_{yy}(0)$). Since $r_{uy}(k)$ is not generally equal to $r_{uy}(-k)$, the cross-correlation function is not symmetric about $k = 0$, but in fact, the cross-correlation coefficients generally tend to zero in the range $(-\infty$ to $i)$ or $(i$ to $+\infty$).

To have a crude check on whether the cross-correlations $r_{uy}(k)$ are effectively zero beyond a certain lag, the cross-correlations are usually plotted out with their 95% confidence intervals, calculated based on the assumption that $u_t$ and $y_t$ are uncorrelated.

Upon this assumption, the variance of the estimated $r_{uy}(k)$ is given by [75]

$$var[r_{uy}(k)] \approx \frac{1}{N - k} \sum_{l=-\infty}^{l=+\infty} r_{uu}(l)r_{yy}(l)$$

(A.300)
Appendix B

Optimal Solution via Analytical Quadratic Programming

B.1 Constraints

For the special case of $NU = 1$, the constraints given in Equation 6.247 can be rewritten as follows:

\[
\begin{align*}
u_{1 \text{min}} & \leq u_{1,t} \leq u_{1 \text{max}} \\
u_{3 \text{min}} & \leq u_{3,t} \leq u_{3 \text{max}} \\
\end{align*}
\]

\[
\begin{align*}
|\Delta u_{1,t}| & \leq \overline{\Delta u_1} \\
|\Delta u_{2,t}| & \leq \overline{\Delta u_2} \\
|\Delta u_{3,t}| & \leq \overline{\Delta u_3} \\
\end{align*}
\]

\[
\begin{align*}
y_{1 \text{min}} & \leq y_{1,t+j} \leq y_{1 \text{max}} \\
y_{2 \text{min}} & \leq y_{2,t+j} \leq y_{2 \text{max}} \\
\end{align*}
\]

(B.301)

The rate constraints on the control signals in Equation B.301 can be modified in a form as:

\[
\begin{align*}
-\overline{\Delta u_1} & \leq \Delta u_{1,t} \leq \overline{\Delta u_1} \\
-\overline{\Delta u_2} & \leq \Delta u_{2,t} \leq \overline{\Delta u_2} \\
-\overline{\Delta u_3} & \leq \Delta u_{3,t} \leq \overline{\Delta u_3} \\
\end{align*}
\]

(B.304)

Since

\[
\Delta u_{2,t} = u_{2,t} - u_{2,t-1} = (u_{2,t}^* + \nabla u_{2,t}) - (u_{2,t-1}^* + \nabla u_{2,t-1})
\]

(B.305)
the rate constraint on \( u_2 \) is modified in a form of the control movement yet to be determined:

\[
-\Delta u_2 - \Delta u_2^* + \nabla u_{2,t-1} \leq \nabla u_{2,t} \leq \Delta u_2 - \Delta u_2^* + \nabla u_{2,t-1} \tag{B.306}
\]

The amplitude constraints on the outputs in Equation B.301 can be written over the predictive horizon \( N_2 \) as follows:

\[
y_{1 \min} = \begin{bmatrix} y_{1 \min} \\ y_{1 \min} \\ \vdots \\ y_{1 \min} \end{bmatrix} \leq \begin{bmatrix} y_{1,t+1} \\ y_{1,t+2} \\ \vdots \\ y_{1,t+N_2} \end{bmatrix} \leq \begin{bmatrix} y_{1 \max} \\ y_{1 \max} \\ \vdots \\ y_{1 \max} \end{bmatrix} = y_{1 \max} \tag{B.307}
\]

and

\[
y_{2 \min} = \begin{bmatrix} y_{2 \min} \\ y_{2 \min} \\ \vdots \\ y_{2 \min} \end{bmatrix} \leq \begin{bmatrix} y_{2,t+1} \\ y_{2,t+2} \\ \vdots \\ y_{2,t+N_2} \end{bmatrix} \leq \begin{bmatrix} y_{2 \max} \\ y_{2 \max} \\ \vdots \\ y_{2 \max} \end{bmatrix} = y_{2 \max} \tag{B.308}
\]

Since

\[
y_1 = G_1' \tilde{u}_1 + G_2' \tilde{u}_2 + f_1 \tag{B.309}
\]

\[
y_2 = G_3' \tilde{u}_2 + G_4' \tilde{u}_3 + f_2
\]

where \( y = [y_{t+1} \ y_{t+2} \ \cdots \ y_{t+N_2}]^T \). Substituting Equation B.309 into B.307 and B.308 gives:

\[
y_{1 \min} - f_1 \leq G_1' \tilde{u}_1 + G_2' \tilde{u}_2 \leq y_{1 \max} - f_1 \tag{B.310}
\]

\[
y_{2 \min} - f_2 \leq G_3' \tilde{u}_2 + G_4' \tilde{u}_3 \leq y_{2 \max} - f_2
\]
The above equation can be modified as follows:

\[
\begin{pmatrix}
\mathbf{y}_{1\text{min}} - f_1
\end{pmatrix} \leq \begin{bmatrix}
G_1' \\
G_2'
\end{bmatrix}
\begin{bmatrix}
\Delta u_{1,t} \\
\nabla u_{2,t} \\
\Delta u_{3,t}
\end{bmatrix}
\leq \begin{pmatrix}
\mathbf{y}_{1\text{max}} - f_1
\end{pmatrix}
\]

(B.311)

The above equation can be modified in a more compact form as:

\[
\begin{pmatrix}
\mathbf{y}_{1\text{min}} - f_1 \\
\mathbf{y}_{2\text{min}} - f_2
\end{pmatrix} \leq \begin{bmatrix}
G_1' \\
0 \\
G_3' \\
G_4'
\end{bmatrix}
\begin{bmatrix}
\Delta u_{1,t} \\
\nabla u_{2,t} \\
\Delta u_{3,t}
\end{bmatrix}
\leq \begin{pmatrix}
\mathbf{y}_{1\text{max}} - f_1 \\
\mathbf{y}_{2\text{max}} - f_2
\end{pmatrix}
\]

(B.312)

Combining Equations B.303, B.304 and B.306 yields:

\[
\begin{align*}
\alpha_1 & \leq \Delta u_{1,t} \leq \beta_1 \\
\alpha_2 & \leq \nabla u_{2,t} \leq \beta_2 \\
\alpha_3 & \leq \Delta u_{3,t} \leq \beta_3
\end{align*}
\]

where

\[
\begin{align*}
\alpha_1 &= \max(-\overline{\Delta u_1}, u_{1\text{min}} - u_{1,t-1}) \\
\beta_1 &= \min(\overline{\Delta u_1}, u_{1\text{max}} - u_{1,t-1}) \\
\alpha_2 &= -\overline{\Delta u_2} - \Delta u_{2,t}^* + \nabla u_{2,t-1} \\
\beta_2 &= \overline{\Delta u_2} - \Delta u_{2,t}^* + \nabla u_{2,t-1} \\
\alpha_3 &= \max(-\overline{\Delta u_3}, u_{3\text{min}} - u_{3,t-1}) \\
\beta_3 &= \min(\overline{\Delta u_3}, u_{3\text{max}} - u_{3,t-1})
\end{align*}
\]

(B.314)

B.2 Optimal Solution via Analytical QP

The following is the derivation of the optimal solutions for the constrained control of the primary refiner for $NU = 1$. The 1st and 2nd order conditions to be satisfied are
Appendix B. Optimal Solution via Analytical Quadratic Programming

rewritten in the follows.

1. The first order conditions are:
\[\frac{\partial J}{\partial \triangle u_{1,t}} = 2h_{11} \triangle u_{1,t} + (h_{12} + h_{21}) \nabla u_{2,t} + 2c_1 + \mu_1 - \mu_2 = 0\]
\[\frac{\partial J}{\partial \nabla u_{2,t}} = (h_{12} + h_{21}) \triangle u_{1,t} + 2h_{22} \nabla u_{2,t} + (h_{23} + h_{32}) \triangle u_{3,t} + 2c_2 + \mu_5 - \mu_6 = 0\]
\[\frac{\partial J}{\partial \triangle u_{3,t}} = (h_{23} + h_{32}) \nabla u_{2,t} + 2h_{33} \triangle u_{3,t} + 2c_3 + \mu_3 - \mu_4 = 0\]  \hspace{1cm} (B.315)

2. The second order conditions are:
\[a_i \mu_i = 0 \text{ for all } i = 1 \text{ to } 6\]
\[\text{such that}\]
\[a_i = 0 \text{ if } \mu_i < 0\] \hspace{1cm} (B.316)
\[\text{or}\]
\[a_i > 0 \text{ if } \mu_i = 0\]
The possible cases of rate and amplitude constraints are given as follows:

Case 1: if \(a_1\) is violated, then
\[\alpha_1 = \triangle u_{1,t}\]
\[\nabla u_{2,t} = \frac{2h_{33}[(h_{12} + h_{21})\alpha_1 + 2c_2] - 2c_3(h_{23} + h_{32})}{(h_{23} + h_{32})^2 - 4h_{23}h_{33}}\]
\[\triangle u_{3,t} = -\frac{(h_{23} + h_{32})}{2h_{33}} \nabla u_{2,t} - \frac{c_3}{h_{33}}\]

Sufficient conditions are:
\[a_i \geq 0, i = 2 \text{ to } 6\]
\[\mu_1 \leq 0\]  \hspace{1cm} (B.317)

Case 2: if \(a_2\) is violated, then
\[\beta_1 = \triangle u_{1,t}\]
\[\nabla u_{2,t} = \frac{2h_{33}[(h_{12} + h_{21})\beta_1 + 2c_2] - 2c_3(h_{23} + h_{32})}{(h_{23} + h_{32})^2 - 4h_{23}h_{33}}\]
\[\triangle u_{3,t} = -\frac{(h_{23} + h_{32})}{2h_{33}} \nabla u_{2,t} - \frac{c_3}{h_{33}}\]
Sufficient conditions are:

\[ a_i \geq 0, i = 1, 3, 4, 5, 6 \]
\[ \mu_2 \leq 0 \]  \hspace{1cm} (B.318)

**Case 3:** if \( a_3 \) is violated, then

\[ \Delta u_{1,t} = -\frac{(h_{12} + h_{21})\alpha_2 + 2c_1}{2h_{11}} \]
\[ \nabla u_{2,t} = \alpha_2 \]
\[ \Delta u_{3,t} = -\frac{(h_{23} + h_{32})\alpha_2 + 2c_3}{2h_{33}} \]

Sufficient conditions are:

\[ a_i \geq 0, i = 1, 2, 4, 5, 6 \]
\[ \mu_3 \leq 0 \]  \hspace{1cm} (B.319)

**Case 4:** if \( a_4 \) is violated, then

\[ \Delta u_{1,t} = -\frac{(h_{12} + h_{21})\beta_2 + 2c_1}{2h_{11}} \]
\[ \nabla u_{2,t} = \beta_2 \]
\[ \Delta u_{3,t} = -\frac{(h_{23} + h_{32})\beta_2 + 2c_3}{2h_{33}} \]

Sufficient conditions are:

\[ a_i \geq 0, i = 1, 2, 3, 5, 6 \]
\[ \mu_4 \leq 0 \]  \hspace{1cm} (B.320)

**Case 5:** if \( a_5 \) is violated, then

\[ \Delta u_{1,t} = \frac{(h_{12} + h_{21})[(h_{23} + h_{32})\alpha_3 + 2c_2] - 4c_1 h_{22}}{4h_{11}h_{22} - (h_{12} + h_{21})^2} \]
\[ \nabla u_{2,t} = -\frac{2h_{11}}{h_{12} + h_{21}} \Delta u_{1,t} - \frac{2c_1}{h_{12} + h_{21}} \]
\[ \Delta u_{3,t} = \alpha_3 \]
Appendix B. Optimal Solution via Analytical Quadratic Programming

Sufficient conditions are:

\[ a_i \geq 0, \quad i = 1, 2, 3, 4, 6 \]

\[ \mu_5 \leq 0 \]  

(B.321)

**Case 6:** if \( a_6 \) is violated, then

\[ \triangle u_{1,t} = \frac{(h_{12} + h_{21})[(h_{23} + h_{32})\beta_3 + 2c_2] - 4c_1 h_{22}}{4h_{11} h_{22} - (h_{12} + h_{21})^2} \]

\[ \nabla u_{2,t} = - \frac{2h_{11}}{h_{12} + h_{21}} \triangle u_{1,t} - \frac{2c_1}{h_{12} + h_{21}} \]

\[ \triangle u_{3,t} = \beta_3 \]

Sufficient conditions are:

\[ a_i \geq 0, \quad i = 1 \text{ to } 5 \]

\[ \mu_6 \leq 0 \]  

(B.322)

**Case 7:** if \( a_1 \) and \( a_3 \) are violated, then

\[ \triangle u_{1,t} = \alpha_1 \]

\[ \nabla u_{2,t} = \alpha_2 \]

\[ \triangle u_{3,t} = - \frac{h_{23} + h_{32}}{2h_{33}} \alpha_2 - \frac{c_3}{h_{33}} \]

Sufficient conditions are:

\[ a_i \geq 0, \quad i = 2, 4, 5, 6 \]

\[ \mu_1, \mu_3 \leq 0 \]  

(B.323)

**Case 8:** if \( a_1 \) and \( a_4 \) are violated, then

\[ \triangle u_{1,t} = \alpha_1 \]

\[ \nabla u_{2,t} = \beta_2 \]

\[ \triangle u_{3,t} = - \frac{h_{23} + h_{32}}{2h_{33}} \beta_2 - \frac{c_3}{h_{33}} \]
Appendix B. Optimal Solution via Analytical Quadratic Programming

Sufficient conditions are:

\[ a_i \geq 0, \ i = 2, 3, 5, 6 \]
\[ \mu_1, \mu_4 \leq 0 \]  \hspace{1cm} (B.324)

**Case 9:** if \( a_1 \) and \( a_5 \) are violated, then

\[ \Delta u_{1,t} = \alpha_1 \]
\[ \nabla u_{2,t} = -\frac{(h_{12} + h_{21})\alpha_1 + (h_{23} + h_{31})\alpha_3 + 2c_2}{2h_{22}} \]
\[ \Delta u_{3,t} = \alpha_3 \]

Sufficient conditions are:

\[ a_i \geq 0, \ i = 2, 3, 4, 6 \]
\[ \mu_1, \mu_5 \leq 0 \]  \hspace{1cm} (B.325)

**Case 10:** if \( a_1 \) and \( a_6 \) are violated, then

\[ \Delta u_{1,t} = \alpha_1 \]
\[ \nabla u_{2,t} = -\frac{(h_{12} + h_{21})\alpha_1 + (h_{23} + h_{31})\beta_3 + 2c_2}{2h_{22}} \]
\[ \Delta u_{3,t} = \beta_3 \]

Sufficient conditions are:

\[ a_i \geq 0, \ i = 2, 3, 4, 5 \]
\[ \mu_1, \mu_6 \leq 0 \]  \hspace{1cm} (B.326)

**Case 11:** if \( a_2 \) and \( a_3 \) are violated, then

\[ \Delta u_{1,t} = \beta_1 \]
\[ \nabla u_{2,t} = \alpha_2 \]
\[ \Delta u_{3,t} = -\frac{h_{23} + h_{32}}{2h_{33}}\alpha_2 - \frac{c_3}{h_{33}} \]
Appendix B. Optimal Solution via Analytical Quadratic Programming

Sufficient conditions are:

\[ a_i \geq 0, i = 1, 4, 5, 6 \]
\[ \mu_2, \mu_3 \leq 0 \]  
(B.327)

Case 12: if \( a_2 \) and \( a_4 \) are violated, then

\[
\begin{align*}
\Delta u_{1,t} &= \beta_1 \\
\nabla u_{2,t} &= \beta_2 \\
\Delta u_{3,t} &= -\frac{h_{23} + h_{32}}{2h_{33}} \beta_2 - \frac{c_3}{h_{33}}
\end{align*}
\]

Sufficient conditions are:

\[ a_i \geq 0, i = 1, 3, 5, 6 \]
\[ \mu_2, \mu_4 \leq 0 \]  
(B.328)

Case 13: if \( a_2 \) and \( a_5 \) are violated, then

\[
\begin{align*}
\Delta u_{1,t} &= \beta_1 \\
\nabla u_{2,t} &= -(h_{12} + h_{21}) \beta_1 + (h_{23} + h_{32}) \alpha_3 + 2c_2 \\
\Delta u_{3,t} &= \alpha_3
\end{align*}
\]

Sufficient conditions are:

\[ a_i \geq 0, i = 2, 3, 4, 5 \]
\[ \mu_2, \mu_5 \leq 0 \]  
(B.329)

Case 14: if \( a_2 \) and \( a_6 \) are violated, then
Appendix B. Optimal Solution via Analytical Quadratic Programming

\[ \Delta u_{1,t} = \beta_1 \]
\[ \nabla u_{2,t} = \frac{(h_{12} + h_{21})\beta_1 + (h_{23} + h_{32})\beta_3 + 2c_2}{2h_{22}} \]
\[ \Delta u_{3,t} = \beta_3 \]

Sufficient conditions are:

\[ a_i \geq 0, i = 1, 3, 4, 5 \]
\[ \mu_2, \mu_6 \leq 0 \]

**Case 15:** if \( a_3 \) and \( a_5 \) are violated, then

\[ \Delta u_{1,t} = -\frac{h_{12} + h_{21}}{2h_{11}} \alpha_2 - \frac{c_1}{h_{11}} \]
\[ \nabla u_{2,t} = \alpha_2 \]
\[ \Delta u_{3,t} = \alpha_3 \]

Sufficient conditions are:

\[ a_i \geq 0, i = 1, 2, 4, 6 \]
\[ \mu_3, \mu_5 \leq 0 \]

**Case 16:** if \( a_3 \) and \( a_6 \) are violated, then

\[ \Delta u_{1,t} = -\frac{h_{12} + h_{21}}{2h_{11}} \alpha_2 - \frac{c_1}{h_{11}} \]
\[ \nabla u_{2,t} = \alpha_2 \]
\[ \Delta u_{3,t} = \beta_3 \]

Sufficient conditions are:
\( a_i \geq 0, i = 1, 2, 4, 5 \)
\( \mu_4, \mu_5 \leq 0 \)

(B.332)

**Case 17:** if \( a_4 \) and \( a_5 \) are violated, then

\[
\triangle u_{1,t} = -\frac{h_{12} + h_{21}}{2h_{11}} \beta_2 - \frac{c_1}{h_{11}} \\
\nabla u_{2,t} = \beta_2 \\
\triangle u_{3,t} = \alpha_3
\]

Sufficient conditions are:

\( a_i \geq 0, i = 1, 2, 3, 6 \)
\( \mu_4, \mu_5 \leq 0 \)

(B.333)

**Case 18:** if \( a_4 \) and \( a_6 \) are violated, then

\[
\triangle u_{1,t} = -\frac{h_{12} + h_{21}}{2h_{11}} \beta_2 - \frac{c_1}{h_{11}} \\
\nabla u_{2,t} = \beta_2 \\
\triangle u_{3,t} = \beta_3
\]

Sufficient conditions are:

\( a_i \geq 0, i = 1 \text{ to } 5 \)
\( \mu_4, \mu_5 \leq 0 \)

(B.334)
Appendix C

Constraint Mapping

The following describes how to map the typical constraints for a SISO system to the expression in the form of future control increments.

The amplitude constraints on the control signal over the control horizon $Nu$:

\[ u_{\text{min}} \leq u(t) \leq u_{\text{max}} \]
\[ u_{\text{min}} \leq u(t + 1) \leq u_{\text{max}} \]
\[ \vdots \]
\[ u_{\text{min}} \leq u(t + Nu - 1) \leq u_{\text{max}} \]

According to $\Delta u(t) = u(t) - u(t - 1)$, above equation can be modified as:

\[ u_{\text{min}} \leq u(t) = \Delta u(t) + u(t - 1) \leq u_{\text{max}} \]
\[ u_{\text{min}} \leq u(t + 1) = \Delta u(t + 1) + \Delta u(t) + u(t - 1) \leq u_{\text{max}} \]
\[ \vdots \]
\[ u_{\text{min}} \leq u(t + Nu - 1) = \Delta u(t + Nu - 1) + \Delta u(t + Nu - 1) + \cdots + \Delta u(t) \leq u_{\text{max}} \]  
(C.336)
Appendix C. Constraint Mapping

\[ u_{\text{min}} - u(t-1) \leq \Delta u(t) \leq u_{\text{max}} - u(t-1) \]
\[ u_{\text{min}} - u(t-1) \leq \Delta u(t) + \Delta u(t+1) \leq u_{\text{max}} - u(t-1) \]
\[ \vdots \]
\[ u_{\text{min}} - u(t-1) \leq \Delta u(t) + \Delta u(t+1) + \cdots + \Delta u(t + Nu - 1) \leq u_{\text{max}} - u(t-1) \]

(C.337)

Above equation can be rewritten in a more compact form as:

\[ u_{\text{min}} - A_2 u(t-1) \leq A_1 \bar{u} \leq u_{\text{max}} - A_2 u(t-1) \]

(C.338)

or

\[ A_1 \bar{u} \leq u_{\text{max}} - A_2 u(t-1) \]
\[ -A_1 \bar{u} \leq -u_{\text{min}} + A_2 u(t-1) \]

(C.339)

where

\[ A_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix} \]

(C.340)

\[ A_2 = [1 \ 1 \cdots 1]^T \]

and \( \bar{u} = [u(t) \ u(t+1) \cdots u(t + Nu - 1)]^T \), \( u_{\text{max}} = [u_{\text{max}} \ u_{\text{max}} \cdots u_{\text{max}}]^T \) and \( u_{\text{min}} = [u_{\text{min}} \ u_{\text{min}} \cdots u_{\text{min}}]^T \).

The rate constraints on the control signal over the control horizon \( Nu \):
Appendix C. Constraint Mapping

\[ \Delta u_{\min} \leq \Delta u(t) \leq \Delta u_{\max} \]
\[ \Delta u_{\min} \leq \Delta u(t + 1) \leq \Delta u_{\max} \]
\[ \vdots \]
\[ \Delta u_{\min} \leq \Delta u(t + Nu - 1) \leq \Delta u_{\max} \]

Above equation can be written in a more compact form as:

\[ \Delta u_{\min} \leq I_{Nu \times Nu} \hat{u} \leq \Delta u_{\max} \]  \hspace{1cm} (C.342)

or

\[ I_{Nu \times Nu} \hat{u} \leq \Delta u_{\max} \]
\[ -I_{Nu \times Nu} \hat{u} \leq -\Delta u_{\min} \]  \hspace{1cm} (C.343)

where \(\Delta u_{\min} = [\Delta u_{\min} \ \Delta u_{\min} \cdots \Delta u_{\min}]^T\) and \(\Delta u_{\max} = [\Delta u_{\max} \ \Delta u_{\max} \cdots \Delta u_{\max}]^T\).

The amplitude constraints on the output over the predictive horizon \(N_2\):

\[ y_{\min} \leq y(t + N1) \leq y_{\max} \]
\[ y_{\min} \leq y(t + N1 + 1) \leq y_{\max} \]
\[ \vdots \]
\[ y_{\min} \leq y(t + N2) \leq y_{\max} \]  \hspace{1cm} (C.344)

Since \(y = G'\hat{u} + f\) where \(y = [y(t+1) \ y(t+2) \cdots y(t+N2)]^T\), above inequality equations can be modified to:

\[
\begin{bmatrix}
  y_{\min} \\
  y_{\min} \\
  \vdots \\
  y_{\min}
\end{bmatrix}
\leq
\begin{bmatrix}
  \text{leq}\ G'\hat{u} \ + \ f \\
  \vdots
\end{bmatrix}
\leq
\begin{bmatrix}
  y_{\max} \\
  y_{\max}' \\
  \vdots \\
  y_{\max}
\end{bmatrix} \hspace{1cm} (C.345)
Above equation can be rewritten in a more compact form as:

\[ y_{\text{min}} \leq G'\tilde{u} + f \leq y_{\text{max}} \]  \hfill (C.346)

or

\[ G'\tilde{u} \leq y_{\text{max}} - f \]
\[ -G'\tilde{u} \leq -y_{\text{max}} + f \]  \hfill (C.347)

where \( y_{\text{min}} = [y_{\text{min}} \cdot \ldots \cdot y_{\text{min}}]^T \) and \( y_{\text{max}} = [y_{\text{max}} \cdot \ldots \cdot y_{\text{max}}]^T \).

Combining Equation C.339, C.343 and C.347 gives:

\[
\begin{bmatrix}
A_1 & \quad -A_1 & \\
I_{Nu \times Nu} & \quad -I_{Nu \times Nu} & \\
G' & \quad -G'
\end{bmatrix}
\begin{bmatrix}
\tilde{u}
\end{bmatrix}
\leq
\begin{bmatrix}
\begin{bmatrix}
A_{\text{max}} - A_2 u(t - 1) \\
-u_{\text{min}} + A_2 u(t - 1)
\end{bmatrix} \\
\begin{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
\triangle u_{\text{max}} \\
-\triangle u_{\text{min}}
\end{bmatrix} \\

y_{\text{min}} - f \\
-y_{\text{max}} + f
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\]  \hfill (C.348)