PROCESS-WIDE PERFORMANCE MONITORING

By

Li Jiang

B. E. (Electrical) Chongqing University, P.R.China
M. E. (Electrical) Chongqing University, P.R.China

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

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

April 1999
© Li Jiang, 1999
Abstract

Today's process industries have shown a great concern on how to improve their product quality. The product quality, however, can be improved only when the process performance has been improved. Process performance monitoring, using only the routine operating data without interfering with the normal process operation, makes it possible to improve process performance and hence is useful to process industries.

In this thesis, a hierarchical performance monitoring system is proposed and tested. The hierarchical monitoring system consists of two levels. The higher-level takes advantage of the advanced statistical regression analysis methods principal component analysis (PCA) and partial least squares (PLS) to assess large amounts of correlated process data. This level is able to provide overall process monitoring and a reliable detection for process abnormality. The lower-level is loop-oriented, and is designed to give detailed performance monitoring and fault diagnosis. It detects loop oscillation and locates the source of the oscillation; it detects high-friction in a valve and evaluates the controller itself. Spectral analysis and time series analysis methods and an adaptive nonlinear modeller (ANM) are used for the purpose of diagnosis at the lower-level. Considering the practical needs, a model-free controller tuning algorithm, iterative feedback tuning (IFT), is also built into the lower-level. Integrated in a complementary manner, the two levels can monitor the process performance with enhanced strength.
Acknowledgments

I would like to take this opportunity to express my thanks to my supervisors, Dr. Guy A. Dumont and Dr. Michael S. Davies, for both the financial assistantship and the academic guidance they provided during the whole period of my Master's program at the University of British Columbia.

I also appreciate all the staff and secretaries of the Pulp and Paper Centre and the secretaries of the Department of Electrical and Computer Engineering for their services and assistance.

Finally, I would like to deliver my thanks to all my colleagues and friends in the Pulp and Paper Centre for their friendship and help.
Table of Contents

Abstract .......................................................... ii

Acknowledgments .................................................. iii

List of Tables ...................................................... vii

List of Figures ..................................................... ix

1 Introduction ..................................................... 1
  1.1 Introduction .................................................. 1
  1.2 Literature Review .......................................... 3
  1.3 Motivation of the thesis .................................. 6
  1.4 Contributions of the thesis .............................. 7
  1.5 Outline of the thesis ...................................... 8

2 Performance Monitoring Using Principal Component Analysis and Partial Least Squares ........................................ 9
  2.1 Introduction .................................................. 9
  2.2 Performance Monitoring Using Principal Component Analysis .......................................................... 10
    2.2.1 PCA Modelling .......................................... 11
    2.2.2 Detection of Process Abnormality Using PCA ........ 13
  2.3 Performance Monitoring Using Partial Least Squares .......... 15
    2.3.1 PLS Modelling .......................................... 16
    2.3.2 Detection of Process Abnormality Using PLS .......... 18
5.3.3 Case 3: Performance degradation caused by de-tuned controller . 75

5.4 Summary ............................................................................. 83

6 Conclusion and Future Work .......................................................... 84

6.1 Conclusion ........................................................................... 84
6.2 Future Work ......................................................................... 85

Bibliography .............................................................................. 87
# List of Tables

3.1 Value Table for Normal Valve Input-Output Relationship ................. 28
3.2 Loop Performance Statistics For Case 1 ...................................... 35
3.3 Detuned PID Controller Parameters .......................................... 37
3.4 Loop Performance Statistics For Case 2 ...................................... 37

4.5 One DOF PID Controller Parameters For Regulatory Control ............. 46
4.6 One DOF PID Controller Parameters For Setpoint Tracking ............... 47
4.7 Two DOF PID Controller Parameters For Regulatory Control .............. 48
4.8 Two DOF PID Controller Parameters For Setpoint Tracking ............... 49

5.9 Upward Messages Format ....................................................... 55
5.10 Percent Variance Captured by PCA ........................................ 59
5.11 Loop Statistics For Case 1 ................................................... 64
5.12 Upward Messages From Loop 1 For Case 1 .................................. 66
5.13 Upward Messages From Loop 3 For Case 1 .................................. 67
5.14 Upward Messages From Loop 2 For Case 1 .................................. 67
5.15 Well and Poorly Tuned Parameters of Loop 3 PID Controller ............ 68
5.16 Loop Statistics For Case 2 ................................................... 71
5.17 Oscillation Index ............................................................. 73
5.18 Upward Messages From Loop 1 For Case 2 .................................. 74
5.19 Upward Messages From Loop 2 For Case 2 .................................. 75
5.20 Upward Messages From Loop 3 For Case 2 .................................. 75
5.21 Original and De-tuned Parameters of Loop 1 PID Controller ............. 75
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.22</td>
<td>Loop Statistics For Case 3</td>
<td>78</td>
</tr>
<tr>
<td>5.23</td>
<td>Upward Messages From Loop 1 For Case 3</td>
<td>79</td>
</tr>
<tr>
<td>5.24</td>
<td>Upward Messages From Loop 2 For Case 3</td>
<td>80</td>
</tr>
<tr>
<td>5.25</td>
<td>Upward Messages From Loop 3 For Case 3</td>
<td>80</td>
</tr>
<tr>
<td>5.26</td>
<td>Parameters of Loop 1 PID Controller Before and After Tuning</td>
<td>80</td>
</tr>
<tr>
<td>5.27</td>
<td>New Loop Statistics For Case 3</td>
<td>83</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Block diagram of a SISO process with control valve ................. 21
3.2 Block diagram of Laguerre network .................................. 24
3.3 Valve nonlinear input-output relationship ............................... 28
3.4 Response and spectrum when valve in normal condition .............. 35
3.5 Response and spectrum when valve in normal condition .............. 36
3.6 Response and spectrum when valve in high friction ................... 37
3.7 Response and spectrum for poorly-tuned controller .................... 38
3.8 Response and spectrum for poorly-tuned controller .................... 39

4.9 Block diagram of the closed-loop system ............................... 41
4.10 Responses of regulatory control using one DOF PID controller before (a) and after (b) tuning ............................................. 46
4.11 Responses of setpoint tracking using one DOF PID controller before (a) and after (b) tuning ............................................. 47
4.12 Responses of regulatory control using two DOF PID controller before (a) and after (b) tuning ............................................. 49
4.13 Responses of setpoint tracking using two DOF PID controller before (a) and after (b) tuning ............................................. 50

5.14 Block diagram of the multi-loop system ............................... 54
5.15 Procedure for higher-level performance monitoring ................... 55
5.16 Procedure for lower-level loop monitoring ............................ 56
5.17 Block diagram of a blend tank level control system ................... 57
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.18</td>
<td>Normal responses of the three loops</td>
<td>58</td>
</tr>
<tr>
<td>5.19</td>
<td>Plot of scores for reference data</td>
<td>60</td>
</tr>
<tr>
<td>5.20</td>
<td>Plot of SPE for reference data</td>
<td>60</td>
</tr>
<tr>
<td>5.21</td>
<td>Plot of $T^2$ for reference data</td>
<td>61</td>
</tr>
<tr>
<td>5.22</td>
<td>Responses of the three loops for case 1</td>
<td>61</td>
</tr>
<tr>
<td>5.23</td>
<td>Plot of new scores for case 1</td>
<td>62</td>
</tr>
<tr>
<td>5.24</td>
<td>Plot of $T^2$ for case 1</td>
<td>63</td>
</tr>
<tr>
<td>5.25</td>
<td>Plot of SPE for case 1</td>
<td>63</td>
</tr>
<tr>
<td>5.26</td>
<td>Spectrum for loop 1: case 1</td>
<td>65</td>
</tr>
<tr>
<td>5.27</td>
<td>Spectrum for loop 3: case 1</td>
<td>65</td>
</tr>
<tr>
<td>5.28</td>
<td>Valve input-output relationship: estimated (dash line) vs. reference (solid line): case 1</td>
<td>66</td>
</tr>
<tr>
<td>5.29</td>
<td>Responses of the three loops for case 2</td>
<td>68</td>
</tr>
<tr>
<td>5.30</td>
<td>Plot of new scores for case 2</td>
<td>69</td>
</tr>
<tr>
<td>5.31</td>
<td>Plot of $T^2$ for case 2</td>
<td>70</td>
</tr>
<tr>
<td>5.32</td>
<td>Plot of SPE for case 2</td>
<td>70</td>
</tr>
<tr>
<td>5.33</td>
<td>Spectrum for loop 1: case 2</td>
<td>71</td>
</tr>
<tr>
<td>5.34</td>
<td>Spectrum for loop 2: case 2</td>
<td>72</td>
</tr>
<tr>
<td>5.35</td>
<td>Spectrum for loop 3: case 2</td>
<td>72</td>
</tr>
<tr>
<td>5.36</td>
<td>Valve input-output relationship: estimated (dash line) vs. reference (solid line): case 2</td>
<td>73</td>
</tr>
<tr>
<td>5.37</td>
<td>Valve input-output relationship: estimated (dash line) vs. reference (solid line): case 2</td>
<td>74</td>
</tr>
<tr>
<td>5.38</td>
<td>Responses of the three loops for case 3</td>
<td>76</td>
</tr>
<tr>
<td>5.39</td>
<td>Plot of new scores for case 3</td>
<td>76</td>
</tr>
<tr>
<td>5.40</td>
<td>Plot of SPE for case 3</td>
<td>77</td>
</tr>
</tbody>
</table>
5.41 Plot of $T^2$ for case 3 ........................................... 77
5.42 Spectrum for loop 1: case 3 ........................................... 78
5.43 Spectrum for loop 3: case 3 ........................................... 79
5.44 Responses after tuning for case 3 ........................................... 81
5.45 Plot of new scores for case 3 ........................................... 81
5.46 Plot of new $SPE$ for case 3 ........................................... 82
5.47 Plot of new $T^2$ for case 3 ........................................... 82
Chapter 1

Introduction

1.1 Introduction

For a long time, the design of advanced control algorithms has been the main preoccupation of the control community. The rationale for this effort has been that there exist a large number of systems in the control field which are complex and difficult to control. So, advanced optimal, non-linear, adaptive or predictive control algorithms are employed to control complex systems. Although today's control engineers have many advanced control techniques available, it is still difficult to find ways to measure or evaluate the control performance effectively once the controller has been installed. The development of techniques for control performance monitoring greatly lags behind the development of control algorithms. However, the situation, in which the research and development have been long neglected, is now changing. The original momentum for this change was generated from the changing world markets, requiring industries to manufacture better quality products with lower cost so as to keep their competitiveness. That means tighter quality control is needed. Generally speaking, the improvement of product quality and the reduction of cost directly results from improving the performance of the manufacturing facilities. The traditional quality control means, such as the statistical process control (SPC) charts represented by the commonly used Shewhart charts, Cusum charts and EWMA charts, are no longer found to be effective-enough quality control measures for modern process industries, although they are still in use. In a typical process plant,
there are hundreds or even thousands of control loops but fairly limited human resources. This makes good system maintenance difficult. Improper tuning of controllers, wear of actuators, malfunction of sensors, and the change of process dynamics often cause the process performance to degrade. According to a report by Bialkowski [4] in 1992, 30% of control loops were oscillating in Canadian pulp and paper mills. This implies that a tough task is waiting for the process industries. Basically, it is impossible to keep a process with such a large number of control loops as mentioned above in normal operating condition without the aid of automated process monitoring and diagnosis tools. On the other hand, if some automated process performance monitoring mechanism is available, the process faults can be detected and located once they occur so that process operators or engineers can be alerted and take corresponding corrective measures immediately. In this way, process performance can be improved and better quality control can be achieved.

In the following section, a brief review of the achievements made in the short history of performance monitoring is given. First, we need to know the guidelines for developing a performance monitoring technique. The two major guidelines are:

1. The performance monitoring technique must be able to detect the degradation of performance and enable the process personnel to take corrective actions in the event of abnormal process behaviours;

2. The performance monitoring technique should avoid introducing an extraneous disturbance or excitation to the process. The application of monitoring is based solely on routine process operating data without interfering with the normal process operation.
1.2 Literature Review

The establishment of the theoretical foundation for modern performance monitoring techniques can be back dated to 1970 when Astrom [3] reported the use of minimum variance control as a benchmark against which to assess control loop performance. In 1978, DeVries and Wu [9] applied the analysis of dispersion and spectral methods to evaluate multivariate process performance. The current effort in performance monitoring started when Harris [20] published his notable work in 1989 in which he showed how to use simple time series analysis techniques to estimate the feedback controller-invariant term from routine operating data of a SISO process with time delay and to assess control loop performance. The contribution of Harris was significant in the sense that it marked a new direction and framework for the loop performance monitoring area. Later on, Desborough and Harris [8, 7, 6] defined a normalized performance index for univariate feedback control and feedback/feedforward control. Lynch and Dumont applied a similar idea to the monitoring of a Kamyr digester chip level control process. Tyler and Morari [46] extended the same idea to assess the performance of unstable and nonminimum-phase SISO processes. Horch and Isaksson [12] suggested an alternative performance index from the standpoint of practical consideration. Huang et al. [25, 24, 26, 23] extended Harris's performance assessment concept to assessing the performance of MIMO feedback controllers.

The techniques based on the Harris's performance assessment concept have become the performance assessment mainstream. Although minimum variance control is seldom used in practice because of its poor robustness and excessive control action, minimum variance control can be used as a benchmark to assess control loop performance by means of Harris's performance index or its extensions. With the performance index at hand, process engineers know whether a loop controller performs well or not. They also know
Introduction

whether further improvement of performance can be achieved by retuning or redesigning the controller, or by modifying the loop structure. To make use Harris's performance index to assess the loop performance, the only \textit{a priori} knowledge is the time delay of the process. While there are various techniques for estimating the process time delay, the variable regression estimation (VRE) algorithm proposed by Elnaggar and Dumont [11] shows an excellent convergence and speed for the delay estimation in open loop. It also has the advantage of estimating delay independently of process parameters. With the knowledge of time delay, the Harris's index can be estimated on-line.

Another important performance monitoring techniques is based on the advanced multivariate statistical analysis methods, principal component analysis (PCA) and partial least squares (PLS). The rationale for use of multivariate statistical analysis methods in process performance monitoring is that the abundant operating data can be used to distill useful information about the process running status. Multivariate statistical analysis methods are superior to their univariate counterparts in handling correlated process data. The PCA and PLS methods are exceptionally suitable for this job. Besides having the properties that other multivariate statistical analysis methods possess, PCA and PLS have the unique property of reducing the dimensionality of problem. The properties of PCA and PLS have attracted the attention of many researchers from different academic fields. Geladi and Kowalski [15] used the PLS method for analysing simulated chemical data in 1986. Hoskuldsson [22] gave a detailed description of using PLS for model building, an important aspect of data analysis. Wangen [48] used a multiblock PLS algorithm for investigating complex chemical systems. Meglen [35] employed PCA for examining large chemical databases. In recent years, PCA and PLS have been applied in process performance monitoring. MacGregor et al. [30, 34, 38, 39] have made a noteworthy contribution in this area. There have been many reports of successful use of PCA or
PLS for process performance monitoring, for example the application of PLS in industrial fed-batch fermentation monitoring by Gregerson [13] and the application of PCA in monitoring emulsion batch processes by Neogi [37]. Other reports of applications of PCA and PLS in process performance monitoring include [1, 10, 36, 42, 45, 47].

While there exist various forms of PCA- or PLS-based techniques, the central idea is the same. First, a reference model representing normal process operation must be built from a set of “good” process data using PCA or PLS. The reference model defines the boundary for good process performance. Then the run-time process behaviour is compared against this reference model. If the process behaviour does not exceed the boundary, then the process is normal or “in-control”; otherwise, it is abnormal or “out-of-control”. After detecting the occurrence of process abnormality, the PCA- or PLS-based procedure will go to identify the variables being responsible for the process abnormality.

In addition to the above-mentioned mainstreams of modern performance monitoring techniques, there are still other kinds of techniques developed for the purpose of sensor auditing, controller and actuator evaluation. Kammer, et al. [29] presented a model-free approach for evaluating the controller in a SISO linear time-invariant process in the sense of LQ optimality. To apply this approach, an exogenous signal is added into the closed-loop process in order to make the measurable signals more informative. By checking the cross-spectral densities obtained from the closed-loop data, the control optimality can be tested. This approach does not need the model of process dynamics, however, it needs an extraneous excitation to the process. Huang [23] proposed the use of LQG benchmark to assess loop performance. This LQG benchmark is more practical than minimum variance benchmark because it takes the control constraint into consideration, but its acquisition needs closed-loop identification. Taha and Dumont [40] investigated the factors causing loop oscillation and described a procedure to detect and diagnose the oscillation via valve checking and controller evaluation. Hagglund [19, 18] proposed a disturbance supervision
scheme for feedback loop monitoring. In this scheme, he addressed the sources of distur-
bances and the factors which could lead to loop oscillation, and developed a procedure
to detect and diagnose the loop oscillation. Taylor [44] investigated in more details the
effect of valve backlash/stiction on loop performance and suggested procedures to specify
and verify control valve performance so as to assure the minimal process variability.

As to the field trials of performance monitoring techniques, Fu and Dumont [14]
developed an on-line control loop performance evaluation algorithm which was tested in
a Canadian pulp and paper mill. Stanfelj, et al. [43] presented a hierarchical system
for monitoring and diagnosing the performance of single-loop control systems. This
system consists of four levels. Each level is designed to implement the detection and
diagnosis task in a top-down mode. It was reported to be used for monitoring a industrial
heat exchanger. Jofriet et al. [28] developed an expert system named QCLiP and had
it to be installed at QUNO Corporation's Thorold Mill. This expert system is DCS-
based and works at a supervisory level with a user-friendly interface. It can evaluate
loop performance based on certain rules and exception report cases. Owen, et al. [41]
implemented a prototype on-line automatic monitoring system at a paper mill. Using
only a small amount of prior information about the loop such as the time delay, the
system can assess and diagnose a number of control loops, based on the data obtained
from a DCS.

There are now commercial performance monitoring packages, which analyse the plant
operating data trend and audit the control loop performance.

1.3 Motivation of the thesis

Although the number of performance monitoring techniques has increased, the range
of their applications is still very limited. Most are variants of the Harris’s index or
Introduction

PCA/PLS analysis methods, which make them entirely not suitable for comprehensive process monitoring. In a typical process, there are hundreds or even thousands of control loops, which in turn consist of even more controllers, actuators and sensors, etc.. What is more, the control loops are often interactive. For such a large scale complex system, a single monitoring technique is incapable of providing effective monitoring in a timely fashion. Therefore, a monitoring system which can provide large-scale process monitoring and diagnosing is required. The system must be able to handle the overwhelming amount of correlated raw process operating data and to provide overall process monitoring using the distilled information. The overall process monitoring allows process personnel to have real-time information of the status of their process so that they can respond to a process fault immediately. On the other hand, the monitoring system is required to provide a detailed and effective fault diagnosing ability. This ability can help the process personnel know where the problem is and who should take the responsibility for early and corresponding corrective action can be taken. To investigate and develop such kind of performance monitoring system becomes the main motivation for this thesis.

1.4 Contributions of the thesis

1. It proposes and implements a hierarchically architectured two-level process-wide performance monitoring system. The higher-level subsystem is process-oriented and aims at providing overall process monitoring; while the lower-level is loop-oriented and aims at providing detailed diagnosis.

2. It successfully employs principal component analysis (PCA) on process data and provides an on-line monitoring solution using sliding window.

3. It implements a valve input-output relationship estimation algorithm based on an adaptive nonlinear modeller (ANM) to evaluate the valve nonlinearity.
4. It implements an oscillation index to locate oscillation caused by the loop controller in a linear SISO loop.

5. It implements a model-free controller tuning algorithm and embeds it into the loop monitoring system.

1.5 Outline of the thesis

The thesis is outlined in the following way. In Chapter 2, the advantages and mechanisms of the advanced statistical analysis methods, principal component analysis (PCA) and partial least squares (PLS), are introduced. The procedures of how to use the PCA and PLS for process performance monitoring are also described in this chapter. Chapter 3 describes the loop-oriented monitoring techniques, which are used for assessing loop performance, detecting and locating loop oscillation, and evaluating valve as well as controller. In Chapter 4, a model free controller tuning algorithm, iterative feedback tuning (IFT), is introduced and implemented. Last, a hierarchically architectured comprehensive performance monitoring system is proposed in Chapter 5.
Chapter 2

Performance Monitoring Using Principal Component Analysis and Partial Least Squares

2.1 Introduction

For a typical process, large quantities of operating data are collected every few seconds, minutes or hours from a multitude of sensors in a multi-loop process [30] in their row form then data values are too voluminous to be assessed by operating personnel. Without any doubt, these data contain information about the process status and should be utilized for performance monitoring. These data, however, are usually correlated one another and make the traditional Statistic Process Control (SPC) approaches often lead to erroneous decisions. Multivariate statistical analysis methods of Principal Component Analysis (PCA) and Partial Least Squares (PLS) can be used to reduce the dimensions of the original correlated data.

In using PCA or PLS for performance monitoring, it is first necessary to build a reference model and then to compare the process behaviour against this model. If the process behaviour does not exceed the boundary defined by the model, then the process is said to be in “normal operating condition” or “in-control”; otherwise, the process is said to be abnormal or “out-of-control”.

9
2.2 Performance Monitoring Using Principal Component Analysis

PCA is appropriate to be used in the case where the data available is of only one type, either all measurements are of process variables or only the block of quality variables is of interest. Then the data can be arranged in a single matrix, $X$, with the samples as its rows and variables as its columns.

PCA is a procedure used to explain the variance in the matrix $X$. One of the most distinguished advantages that PCA possesses is that it can greatly reduce the dimensionality of the problem when the variables in the data set are highly correlated. When PCA is used, it decomposes the $X$ into a sum of $k$ rank 1 matrices, which are outer products of the vectors called scores and loadings. The first loading vector or the first principal component defines the direction of greatest variability in $X$. The first principal component is in fact the eigenvector of $X^T X$ associated with the largest eigenvalue. The second principal component is orthogonal to the first principal component, and explains the greatest amount of the remaining variability. The same procedure will be repeated until a stopping criterion is satisfied. The stopping criterion is chosen so that most of the variation in the data matrix $X$ has been explained, and the number of principal components is then accordingly determined. Since the variables in the matrix $X$ are correlated, the number of principal components is generally much less than that of the original variables. Therefore, most of the information contained in the original data set can be represented with smaller dimensions. Algebraically, the data matrix $X$ can be decomposed as

$$X = t_1 p_1^T + t_2 p_2^T + ... + t_k p_k^T + E$$

(2.1)

where $p_i$ is the $i$th principal component or loading vector, and $t_i$ is the $i$th score vector($i=1...k$). $k$ is the number of principal components. $E$ is a residual matrix. Ideally,
k is chosen such that there is no significant process information left in \( E \).

There are several stopping criterions for determining this maximum significant dimension \( k \). The procedure of cross-validation [23, 39], however, is a robust method and has been used widely.

### 2.2.1 PCA Modelling

To build a model for performance monitoring, the first step is to select a set of reference data which represent normal process performance. Then, pretreatment will be applied to the reference data. Generally speaking, the data pretreatment includes data scaling, missing data handling, and outlier screening.

For the purpose of performance monitoring, data scaling is carried in the form of either mean-centered scaling or autoscaling. Mean-centered scaling is to scale the data such that each variable has a zero mean. Autoscaling scales the data such that each variable is in standard units, i.e., has zero mean and unit standard deviation. When all the variables are treated to be identically important, autoscaling is used.

The collected data may have variables with missing entries caused by malfunction of sensors, transmission errors, and so on. To fully utilize the existing data, the missing data must be handled properly. The common method is to fill the missing data with the mean value determined from the related variable samples.

Outliers are influential observations, which do not comply to the distribution pattern of the corresponding variables. To establish a reliable model, outliers must be detected and then be eliminated. One of the commonly used methods is to use PCA iteratively.

After the reference data set has been pre-treated, PCA can be employed to build the reference model. Instead of using the standard PCA, which involves solving all the principal components at once, a nonlinear iterative partial least squares (NIPALS) algorithm [16, 15], which is more computation-effective, is used. The NIPALS algorithm
is described as below

i. Scale the $X$

ii. Choose arbitrary column of $X$ as $t$

iii. 

$$E = X$$ \hspace{1cm} (2.2)

1. 

$$P = E^T \cdot t$$ \hspace{1cm} (2.3)

2. 

$$P = P/\|P\|$$ \hspace{1cm} (2.4)

3. 

$$t = E \oslash P$$ \hspace{1cm} (2.5)

4. if $t$ has converged then go to step 5, otherwise go to step 1.

5. 

$$E = E - t \otimes P$$ \hspace{1cm} (2.6)

iv. Go to step 1 to calculate the next principal component until some stopping criterion is satisfied. Here, cross-validation is used as the stopping criterion for the procedure.

When the above procedure stops, information about the principal components, the number, the loading and score vectors, has been obtained. The boundary for the “normal operating condition” can then be defined. This boundary can be graphically represented by an ellipse in the case of only two principal components or can be placed statistically using Hotelling’s $T^2$ statistic which is directly related to the $F$-distribution [27, 34]. Given the number of the principal components and the confidence level, Hotelling’s $T^2$ statistic, which defines the limit of score distance in the new coordinate system, is given as

$$T^2_{k,n,\alpha} = \frac{k(n-1)}{n-k} F_{k,n-k,\alpha}$$ \hspace{1cm} (2.7)
where \( n \) is the number of observations, \( k \) is the number of principal components, and \( \alpha \) is the confidence level.

Another important control limit, the standard prediction error (SPE), should be included in the model. This statistic defines the limit for testing if an unusual pattern of variation, which cannot be explained by the model, occurs. According to Gregerson [13], the \( SPE \) limit is calculated by

\[
SPE_\alpha = (v/2m)\chi^2_{2m^2/v, \alpha}
\]

where \( m \) and \( v \) are the mean and variance of the \( SPE \) sample at each time, \( \alpha \) is the confidence level.

So far, the reference model has been established. The model consists of the number of principal components, their loading and score vectors, the \( T^2 \) statistic and the \( SPE \) limit.

### 2.2.2 Detection of Process Abnormality Using PCA

Once the reference model is established, it can be used to monitor the process performance by detecting process abnormality and identifying the contributing variables.

To detect process abnormality, the first step is to project the new process observation into the new coordinate system defined by the reference model as

\[
t = xP
\]

where \( t \) is the location of the original observation in the new coordinate system, and \( P \) is the matrix of loading vectors.

According to Nomikos and MacGregor et al. [34, 38, 39] and Gregerson [13], the score distance, namely the \( T^2 \) statistic, and the \( SPE \) can be used to detect the process
abnormality. The $T^2$ statistic is calculated by

$$T^2 = t^T\Sigma^{-1}t$$  \hspace{1cm} (2.10)

where $\Sigma$ is the covariance matrix of the independent variables in the new coordinate system. It is a diagonal matrix.

Then, $T^2$ statistic is compared against the nominal $T^2_{k,n-k,\alpha}$ statistic. If $T^2$ is greater than $T^2_{k,n-k,\alpha}$, that means some excessive variation occurs to the process. According to Anderson [2], the boundary for the "normal region" can be defined using an ellipse represented by the following equation.

$$(\bar{t} - m)^T\Sigma^{-1}(\bar{t} - m) \leq T^2_{k,n,\alpha}$$  \hspace{1cm} (2.11)

where $m$ is the new dimension reduced space.

The $SPE$ will also be computed and compared to the control limit. The $SPE$ of the observation $x_i$ is calculated in the following way

$$SPE_{x_i} = \sum_{j=1}^{k}(x_{ij} - \hat{x}_{ij})^2$$  \hspace{1cm} (2.12)

where $\hat{x}_{ij}$ is the prediction of $x_i$ from the reference model.

When the $SPE$ statistic exceeds its limit, that means some new event has occurred in the process and the model is no longer valid.

If the process behaves well, the scores will fall into the ellipse. When the scores fall outside the ellipse, there exists excessive variation of process, which can be explained by the model. The $SPE$ limit is used to check whether any unusual event, which can not be explained by the model, occurs to the process. When both the $T^2$ statistic and the $SPE$ are below their control limits, then the process is in "normal operating condition"; otherwise, some process abnormal event occurs. Once the occurrence of
process abnormality has been detected, the responsible contributing variables should be identified so as to facilitate the further diagnosis. Unfortunately, it is a difficult task to find the reason for the process abnormality. One solution is to develop an expert system based on the behaviour of the data projections in the reduced space. The contributing variables, however, can be identified when the $SPE$ exceeds its control limit. In this case, the responsible variables can be identified by checking the individual contributions of variables to the $SPE$ at the time the process abnormality occurs. The variables which contribute the most will be regarded as the responsible variables.

2.3 Performance Monitoring Using Partial Least Squares

The mechanism of using PLS for process monitoring is similar to that of using PCA. The PLS method, however, has some additional properties making it superior to PCA. In PCA only a single data matrix can be handled. Often there are variables such as product quality or productivity that are of greater importance and should be included in the monitoring procedure. Unfortunately, the quality or productivity variables are often measured on a much less frequent basis than the process variables and so the information contained in the process variables must be used to predict, monitor and detect the changes in the process quality variables. The PLS algorithm provides a good solution to the problem.

Conceptually, there no major difference between PCA and PLS except that PLS can simultaneously reduce the dimensional spaces of the two highly correlated data sets into lower dimensional spaces defined by their latent vectors. Assume that $X$ is the matrix of process variables, and $Y$ is the matrix of quality variables. Algebraically, the data matrices $X$ and $Y$ can be decomposed as
Performance Monitoring Using Principal Component Analysis and Partial Least Squares

\[
X = t_1 p_1^T + t_2 p_2^T + \ldots + t_k p_k^T + E \quad (2.13)
\]

\[
Y = u_1 q_1^T + u_2 q_2^T + \ldots + u_k q_k^T + F \quad (2.14)
\]

where \( p_i \) and \( t_i \) are the \( i \)th latent vector and score vector of matrix \( X \) respectively, and \( E \) is its associated residual matrix. \( q_i \) and \( u_i \) are the \( i \)th latent vector and score vector of matrix \( Y \) respectively, and \( F \) is the associated residual matrix. \( k \) is the number of latent variables.

The regression part of PLS is an inner relationship given by

\[
\hat{u}_i = b_i t_i \quad (2.15)
\]

where \( b_i = u_i^T t_i / t_i^T t_i \). The “hat” indicates that the vector is an estimated one.

This leads to the following relationship

\[
Y = TBQ^T + F \quad (2.16)
\]

where \( TBQ^T \) is the prediction of \( Y \).

### 2.3.1 PLS Modelling

The first step in PLS modelling is also to select two sets of reference data containing process variables and quality variables, respectively. Data pretreatment is applied to the selected data sets. After the data are well prepared, the iterative PLS algorithm will be used to determine the latent variables. Then, the \( T_k^2 \) statistics and the \( SPE_x \) and \( SPE_y \) limits will be calculated for the definition of “normal operating condition”. The calculation procedures are similar to those used in the PCA case.
To obtain the number of latent variables as well as their loading and score vectors, an iterative algorithm is used. The iterative algorithm is represented by the procedure below.

1. Start: set \( u \) to be an arbitrary column of \( Y \).

2. 

\[
    w^T = u^T X / u^T u \tag{2.17}
\]

3. 

\[
    w^T = w^T / ||w^T|| \tag{2.18}
\]

4. 

\[
    t = Xw / w^T w \tag{2.19}
\]

5. 

\[
    q^T = t^T Y / t^T t \tag{2.20}
\]

6. 

\[
    q^T = q^T / ||q^T|| \tag{2.21}
\]

7. 

\[
    u = Yq / q^T q \tag{2.22}
\]

8. Check convergence: if yes go to step 9, otherwise go to step 2.

9. 

\[
    p = X^T t / t^T t \tag{2.23}
\]

10. 

\[
    b = u^T t / t^T t \tag{2.24}
\]

11. 

\[
    E = X - tp^T \tag{2.25}
\]
and

$$F = Y - btq^T$$ \hspace{1cm} (2.26)

12. Go to step 2 to calculate the next latent variable until some stopping criterion is satisfied.

The cross-validation is used as the stopping criterion.

2.3.2 Detection of Process Abnormality Using PLS

The procedure used to detect process abnormality is similar to that used by PCA. The first step is also project the new observation of process variables to the new coordinate system. Then, the $T^2$ statistic will be computed and compared with the nominal one. Additionally, the $SPEs$ of $X$ and $Y$ will also be calculated and tested to see whether they are under the control limits or not.

The $T^2$ statistic is calculated by

$$T^2 = t^T \Sigma^{-1} t$$ \hspace{1cm} (2.27)

where $\Sigma$ is the covariance matrix of the independent process variables in the new coordinate system. It is a diagonal matrix.

And the following equations are used to calculate $SPE_x$ and $SPE_y$

$$SPE_{x,i} = \sum_{j=1}^{k} (x_{ij} - \hat{x}_{ij})^2$$ \hspace{1cm} (2.28)

$$SPE_{y,i} = \sum_{j=1}^{k} (y_{ij} - \hat{y}_{ij})^2$$ \hspace{1cm} (2.29)

where the variables with “hats” are predicted ones.

If the $T^2$ statistic and the $SPEs$ are below the corresponding limits, the process is normal; otherwise, the process is abnormal. If only the $T^2$ statistic exceeds the limit, the
variation of the process is too great and the model is valid. If the $SPE$ statistic exceeds its limit, then some new event has occurred and the model is no longer valid.

The same difficulty in identifying the variables responsible for the process abnormality exists as that for PCA. However, the responsible contributing variables can be identified by checking the individual contributions of process variables to the $SPE$ if the process abnormality is featured by excessive $SPE$.

2.4 Summary

The multivariate statistical analysis methods Principal Component Analysis and Partial Least Squares are very suitable for monitoring a process where highly correlated data are encountered. To monitor a process, a reference model is first established using a reference set of data collected when the process performance meets the prescribed specification; then the process behaviour is compared against this model. Two kinds of statistics are used to detect the process abnormality. One is the $T^2$ statistic, and the other is $SPE$ statistic. In the case when the process abnormality is detected by finding excessive $SPE$, the responsible variables can be identified out by checking the contributions from the individual variables to the statistics.
Chapter 3

Loop Performance Monitoring

3.1 Introduction

A typical manufacturing process contains many control loops. The performance of the control loops is subject to deterioration due to many factors. Although the use of PCA and PLS can detect occurrence of process abnormality, they are not effective means to diagnose the original causes for the abnormality. Therefore, a monitoring mechanism is needed for diagnosis when loop performance deteriorates. The loop performance monitoring procedure proposed here is oriented to a single-input- single-output control loop, and will be integrated into each control loop to diagnose common causes for the deterioration of a control loop. Since control valves are widely used in process control loops, the control loop used to test the proposed loop performance monitoring procedure includes a control valve, as shown in Figure 3.1.

3.2 Performance Index and Controller Evaluation

The performance of an existing control loop is measured against a benchmark, such as offset from setpoint, overshoot, rise-time, and variance. For regulatory control, process variance is an important performance measure since many quality criteria are based on it. Traditionally, the performance of a control loop might be deemed unacceptable if the output variance exceeds some critical value. This criterion, however, fails to recognize the difference between achievable and acceptable performance. Some cases are discussed
by Harris [20], where the controller is already giving the best possible performance, i.e. it is not possible to reduce the variability of a process variable by simply re-tuning or re-designing a linear loop controller, but the resulting performance is still unacceptable from the perspective of quality specification. In these cases, reductions in variability are only achieved by modifying the system or by employing a nonlinear controller. To recognize the difference between acceptable performance and good control, a performance benchmark is needed. One natural and logical choice is minimum variance control.

Many industrial processes can be adequately modelled in discrete time as

\[ y(t) - \mu = \frac{B(z^{-1})}{A_1(z^{-1})}z^{-k}u(t) + \frac{C(z^{-1})}{A_2(z^{-1})}\nabla d e(t) \]  

(3.30)

where \( y(t) \) is the measured process output, \( \mu \) is the mean of \( y(t) \), and \( u(t) \) is the deviation of manipulated variable from a reference value required to keep the process at its mean value. \( k \) is the process time delay, \( \nabla = 1 - z^{-1} \) and \( \{e(t)\} \) is a sequence of white noise subjected to \( N(0, \sigma_e^2) \) distribution.
Under the linear time-invariant feedback control,

\[ u(t) = -\frac{N(z^{-1})}{D(z^{-1})} y(t) \]  \hspace{1cm} (3.31)

it can be readily shown [20] that the closed-loop system is described by

\[ y(t) - \mu_y = H(z^{-1})e(t) \]  \hspace{1cm} (3.32)

where \( \mu_y \) is the mean of \( y(t) \) under feedback control.

The monic polynomial \( H(z^{-1}) \) can be decomposed into two parts as shown below

\[ y(t) - \mu_y = F(z^{-1})e(t) + z^{-k}G(z^{-1})e(t) \]  \hspace{1cm} (3.33)

where \( F(z^{-1}) \) is a monic polynomial of order \( k - 1 \).

\[ F(z^{-1}) = 1 + f_1 z^{-1} + \cdots + f_{k-1} z^{-k+1} \]  \hspace{1cm} (3.34)

The first and second terms in Equation 3.33 can be interpreted as the \( k - \text{step} \) ahead forecast error and \( k - \text{step} \) ahead forecast, respectively. Since \( F(z^{-1}) \) only depends on \( C(z^{-1}) \) and the time delay \( k \), it is feedback invariant.

When minimum variance control is applied to the system, the output is solely determined by the \( k - \text{step} \) ahead forecast error. The output variance then becomes as

\[ \text{var}\{y(t)\} = \text{var}\{F(z^{-1})e(t)\} = \sigma_{mv}^2 \]  \hspace{1cm} (3.35)

In general situations, minimum variance control is not recommended because of its poor robustness and excessive control action. The obtained minimum variance \( \sigma_{mv}^2 \), however, can be used as the benchmark against which to assess control loop performance. To assess of loop performance, a normalized performance index is defined by Desborough.
and Harris [8]. Based on the consideration of practical applications, the normalized performance index uses mean square error of the process output, instead of its variance, to compare against the minimum variance benchmark. The normalized performance index is represented by

\[
PI(k) = 1 - \frac{\sigma_{mv}^2}{\sigma_y^2 + \mu_y^2}
\]  

(3.36)

where \( PI(k) \) is bounded within \([0, 1]\).

When \( PI(k) = 0 \), the loop is under minimum variance control. When \( PI(k) = 1 \), it implies that no control action is applied.

To calculate the performance index, the minimum variance \( \sigma_{mv}^2 \) must be obtained. The traditional methods to compute the \( \sigma_{mv}^2 \) are based on an ARMA model and so share the disadvantage of having to determine the order of the model [5]. This, however, is not a trivial task. To overcome this drawback, Lynch and Dumont [33] proposed an alternative approach based on Laguerre network, which is adapted here due to its straightforwardness. A block diagram of the Laguerre network which models the closed loop noise filter is shown in Figure 3.2.

The discrete Laguerre filters can be represented by

\[
L_i(q) = \frac{\sqrt{1 - a^2}}{z - a} \left( \frac{1 - az}{z - a} \right)^{i-1}, \quad i = 1, \ldots
\]  

(3.37)

where \( a \) is the Laguerre filter time scale and used as a design parameter.

As the Laguerre functions are orthonormal and complete in \( L_2[0, \infty) \), they can represent the impulse response \( h(t) \) of any stable sampled linear system, with an infinite expansion

\[
h(t) = \sum_{i=1}^{\infty} g_i L_i(t)
\]  

(3.38)
where $g_i$ is the $i$-th Laguerre gain, and $l_i(t)$ is the output of the $i$-th Laguerre filter.

In practice, $N$ Laguerre filters are sufficient to represent the impulse response. The transfer function of the impulse response, $H(z^{-1})$, then can be approximated as

$$
H(z^{-1}) = \sum_{i=1}^{N} g_i L_i(z^{-1})
$$

The design parameter $a$ can be chosen by experience or by trials. Once $a$ and the filter number $N$ are selected, the Laguerre gains that best approximate the impulse response are the only things to be determined. For the sake of convenience, the discrete Laguerre network is represented in state-space form as

$$
l(t + 1) = Al(t) + be(t)
$$

$$
y(t) = c^T l(t) + e(t)
$$
where $A$ is a matrix whose elements are given by:

$$A = \begin{cases} 
    a_{ij} = a & i = j \\
    a_{ij} = 0 & i < j \\
    a_{ij} = (-a)^{i-j-1}(1-a^2) & i > j 
\end{cases}$$

(3.42)

and the $b$ matrix elements are defined as

$$b_i = (-a)^{i-1}\sqrt{1-a^2} \quad i = 1, \ldots, N$$

(3.43)

c is the vector of Laguerre gains, and $l$ is the vector of the Laguerre filter outputs. The two vectors are defined as

$$c = [g_1 \ g_2 \cdots \ g_N]^T$$

(3.44)

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

(3.45)

The minimum variance of output consists of the first $k$ terms of the output $y(t)$, given the unmeasurable noise $\{e\}$ as the input. By solving the above state space equation of Laguerre network, the minimum variance is determined as

$$\sigma_{mv}^2 = \sigma_e^2[1 + (c^Tb)^2 + (c^TAb)^2 + \cdots + (c^TA^{k-2}b)^2]$$

(3.46)

where $k$ is the process time delay.

Since the terms $A^ib$ in the above equation only rely on the Laguerre time scale $a$, they become known once the design parameter $a$ is selected. As a result, only the estimates of $c$ and $\sigma_e^2$ are needed for obtaining the estimate of $\sigma_{mv}^2$. This, however, can be done by performing the recursive extended least squares (RELS) estimation algorithm as described below.
\[ l(t) = Al(t - 1) + b\eta(t - 1) \quad (3.47) \]
\[ P(t) = P(t - 1) - \frac{P(t - 1)l(t)l^T(t)P(t - 1)}{1 + l^T(t)P(t - 1)l(t)} \quad (3.48) \]
\[ \hat{c}(t) = \hat{c}(t - 1) + P(t)l(t)[y(t) - \hat{c}^T(t - 1)l(t)] \quad (3.49) \]
\[ \eta(t) = y(t) - \hat{c}^T(t)l(t) \quad (3.50) \]

The residual \( \eta(t) \) gives an estimate of the white noise \( e(t) \), and can be used to estimate \( \sigma^2 \). After the estimates converge, the estimate of \( \sigma_{mv}^2 \) is computed easily by

\[ \hat{\sigma}_{mv}^2 = \hat{\sigma}_e^2[1 + (\hat{c}^Tb)^2 + (\hat{c}^TAb)^2 + \cdots + (\hat{c}^T A^{k-2}b)^2] \quad (3.51) \]

Refer to [33] for more details of using Laguerre network for minimum variance estimation. Given the estimated \( \hat{\sigma}_{mv}^2 \), the performance index can be calculated and used to assess the loop performance.

### 3.3 Oscillation Detection and Location

Oscillation in control loops is a major cause of degradation of process performance. According to Bialkowski [4], about 30% of control loops in Canadian pulp and paper industries are oscillating. Oscillation has a serious negative effect on production in the forms of degrading the product quality, causing raw material loss, and increasing energy consumption. The detection and elimination of oscillation, therefore, is an important task to a loop performance monitoring system. The diagnosis and location of oscillation, however, may be difficult because of the interaction between the control loops. To facilitate the diagnosis, a good understanding to the major causes of loop oscillation is very helpful.
Based on the experience of control practice in process industries, the major factors causing a loop to oscillate are control valves with high friction, poorly tuned or de-tuned loop controllers, and oscillation imported from the interacted loops or caused by load disturbance. Among the three factors, the first one is believed to be most common. This is because a valve will be worn out when it has been used for a long period. The wear of valve will cause high static friction in it, which in turn will lead to severe stiction/backslash effect on the valve. The valve then becomes a typical nonlinear component in the loop and makes the loop oscillate. The loop controllers may also cause oscillations because they could be improperly tune or the change of dynamics process makes them de-tuned. Oscillation can also be caused by load disturbance or be imported from the other loops because of the loop interaction. The diagnosis of imported oscillation is usually based on the results of valve and control diagnosis. Therefore, an effective way to evaluate valve and controller is necessary for loop performance monitoring.

To detect the oscillation of a control loop, an easy and reliable way is to check the power spectrum of the loop output measurement. If spectral power spike amplitude exceeds a certain prescribed threshold, then the decision that the loop is oscillating can be safely made. Once an oscillation is detected within a control loop, the constituent frequencies corresponding to each spectral power spike can also be found. The constituent frequencies are useful for further oscillation location.

Upon detecting the occurrence of oscillation in the loop, the next task is to locate the source of the oscillation. Oscillation might be generated within the loop or imported from an interacted loop. When the oscillation is generated within the loop, the most likely sources are a control valve with high friction or a poorly or de-tuned tuned controller. Since a control valve is more often the primary trouble-maker [4], it will be checked with higher priority than the controller.
3.3.1 Valve High Friction Check

For an ideal valve, its output will track the valve reference signal given by the controller, and its input-output is linear. In practice, valves are usually nonlinear. The nonlinear input-output relationship is illustrated in Figure 3.3.

![Figure 3.3: Valve nonlinear input-output relationship](chart)

A table, as shown in Table 3.1, with input and output values as its entries can be established for this input-output relationship.

<table>
<thead>
<tr>
<th>$u_i$</th>
<th>$u_1$</th>
<th>$u_2$</th>
<th>$u_3$</th>
<th>$u_4$</th>
<th>$u_5$</th>
<th>$u_6$</th>
<th>$u_7$</th>
<th>$u_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(u_i)$</td>
<td>$f(u_1)$</td>
<td>$f(u_2)$</td>
<td>$f(u_3)$</td>
<td>$f(u_4)$</td>
<td>$f(u_5)$</td>
<td>$f(u_6)$</td>
<td>$f(u_7)$</td>
<td>$f(u_8)$</td>
</tr>
</tbody>
</table>

For a normal valve, its input-output relationship can be approximately regarded as
linear around its operating region although its overall input-output relationship is non-linear. This assumption, however, will not valid when the valve is worn out and the friction in it becomes high. High friction in a valve leads to severe static stiction effect, which causes the valve to track the reference signal with delay and overshoot. At this time, the valve becomes as a typical nonlinear component in the loop, and causes the loop to oscillate. A high friction valve can so be detected by checking the degree of the non-linearity of its input-output relationship.

A useful method to represent the nonlinear input-output relationship is the adaptive nonlinear modeller (ANM) proposed by Astrom [32]. The ANM represents the nonlinear input-output relationship by using the following linear interpolation formula in Equation 3.52.

\[
y = f(u) = \frac{u_{i+1} - u}{u_{i+1} - u_i} f(u_i) + \frac{u - u_i}{u_{i+1} - u_i} f(u_{i+1})
\]

(3.52)

where \( u_i \) and \( f(u_i) \) are the entries in the above table.

For any arbitrary input which lies between the interval \( u_{i+1} > u > u_i \) in the table, the output can be determined by using the interpolation formula. Given the valve inputs and measurements, the values \( f(u_i) \) can be estimated for the table using recursive least squares (RLS) so that the input-output relationship can be establish. To use RLS, the model must be placed into a linear regression form as

\[
y = a_1 f(u_1) + a_2 f(u_2) + \cdots + a_{n-1} f(u_{n-1}) + a_n f(u_n)
\]

(3.53)

Since the input \( u \) lies within only one interval \([ u_i \ u_{i+1} ]\), all but two of the \( a_i \)'s will be zero. Using the linear interpolation formula, the values of the two nonzero \( a_i \)'s are determined as
\[ a_i = \frac{u_{i+1} - u}{u_{i+1} - u_i} \]  
\[ a_{i+1} = \frac{u - u_i}{u_{i+1} - u_i} \]  \hspace{1cm} (3.54) \hspace{1cm} (3.55)

For every sample data, after finding which interval \([u_i, u_{i+1}]\) the input \(u\) falls in, the regression vector and the parameter vector can be defined, respectively, as

\[ \phi(t) = [0 \cdots 0 a_i a_{i+1} 0 \cdots 0]^T \]  \hspace{1cm} (3.56)

\[ \theta = [f(u_1) f(u_2) \cdots f(u_i) f(u_{i+1}) \cdots f(u_n)]^T \]  \hspace{1cm} (3.57)

This gives the linear regression equation as below

\[ y = \phi(t)\theta^T + \nu(t) \]  \hspace{1cm} (3.58)

Then the following RLS can be now applied to estimate \(\theta\).

\[ K(t) = \frac{P(t-1)\phi(t)}{1 + \phi^T(t)P(t-1)\phi(t)} \]  \hspace{1cm} (3.59)

\[ \hat{\theta}(t) = \hat{\theta}(t-1) + K(t)[y(t) - \phi(t)\hat{\theta}^T] \]  \hspace{1cm} (3.60)

\[ P(t) = P(t-1) - K(t)\phi^T(t)P(t-1) \]  \hspace{1cm} (3.61)

For checking the non-linearity of the control valve, a normal input-output relationship must be estimated in advance. When oscillation is detected in the loop, RLS will estimate the actual the sampled data. The estimated relationship will then be compared against
the reference relationship. If the estimated relationship curve falls within the predefined "Normal region," then the valve is regarded as in normal condition; otherwise, high friction is assured. The criterion for the evaluation is given as below

\[ I_{fr} = \frac{D}{D_{max}} \quad (3.62) \]

where valve input-output relationship and reference valve input-output relationship; while \( D \) is the actual distance between the estimated valve input-output relationship and reference valve input-output relationship, corresponding to the same valve input.

If the valve is found to exhibit high friction, then the valve should take the responsibility for the oscillation. To eliminate the oscillation, valve maintenance or replacement is needed.

### 3.3.2 Controller Check

If the control valve is checked to be normal, the controller must then be checked since a poorly tuned or de-tuned controller can also cause oscillation. Due to the fact that the valve is normal, the process can be approximately regarded as a linear system around the operating point. If the oscillation is generated within the loop and its frequency is \( \omega \), then the following equation holds

\[ 1 + C(j\omega)P(j\omega) = 0 \quad (3.63) \]

where \( C(j\omega) \) is the transfer function of controller and \( P(j\omega) \) is the transfer function of process plus the valve of the SISO control loop as shown in Figure 3.1.

Given the error signal \( e \) and the output \( y \), the following two equations hold

\[ \frac{Y}{E} = |C(j\omega)P(j\omega)| = 1 \quad (3.64) \]
where \( E \) and \( Y \) are the amplitudes of \( e \) and \( y \), respectively. and

\[
\phi_e + \phi_y = -\pi \tag{3.65}
\]

The signals \( e(t) \) and \( y(t) \) can be expanded into their Fourier series. Since the process is usually a low-pass filter, the high frequency harmonics of the two signals will be filtered out by the process. The signals then can be approximated as

\[
e(t) \approx e_1 \cos(\omega t) + e_2 \sin(\omega t) \tag{3.66}
\]

\[
y(t) \approx y_1 \cos(\omega t) + y_2 \sin(\omega t) \tag{3.67}
\]

The amplitudes of \( e \) and \( y \) can be calculated by

\[
e_1 = \frac{2}{N} \sum_{i=0}^{N} e(i) \cos(\omega i) \tag{3.68}
\]

\[
e_2 = \frac{2}{N} \sum_{i=0}^{N} e(i) \sin(\omega i) \tag{3.69}
\]

\[
y_1 = \frac{2}{N} \sum_{i=0}^{N} y(i) \cos(\omega i) \tag{3.70}
\]

\[
y_2 = \frac{2}{N} \sum_{i=0}^{N} y(i) \sin(\omega i) \tag{3.71}
\]

where \( N \) is the number of samples within a whole period.

\[
E = \sqrt{e_1^2 + e_2^2} \tag{3.72}
\]

\[
Y = \sqrt{y_1^2 + y_2^2} \tag{3.73}
\]

and the phases by

\[
\phi_e = \arctan\left(\frac{e_1}{e_2}\right) \tag{3.74}
\]
\[ \phi_y = \arctan \left( \frac{y_1}{e_2} \right) \quad (3.75) \]

For the sake of convenience, an internal oscillation index defined by Ettaleb [31]

\[ I_{osc}(\omega) = \left| 1 - \frac{Y}{E} \right| \quad (3.76) \]

The oscillation frequencies obtained in oscillation detection will be used to decide whether the oscillation is generated by the loop. The oscillation is deemed to be generated by the loop if there exists a frequency \( \omega \) that can satisfy the condition

\[ I_{osc}(\omega) = \left| 1 - \frac{Y}{U} \right| \approx 0 \quad (3.77) \]

Using the internal oscillation index, if the oscillation is found to be generated by the loop, it implies that the controller is poorly tuned or de-tuned and causes the oscillation.

### 3.4 Simulation

A flow control loop is simulated to test the fault detection and diagnosis abilities of the loop performance monitoring techniques introduced in the foregoing section. In the simulations, two cases are considered. In the first case, the valve is set a state of high friction and causes the loop oscillation. In the second case, the controller is poorly tuned and so results in an oscillation.

The transfer function of the valve is given as below

\[ V(s) = \frac{4}{1 + 8s}e^{-2s} \quad (3.78) \]

The sampling period is 1 second. After discretizing, the valve dynamics are

\[ V(z^{-1}) = \frac{0.47}{1 - 0.8825z^{-1}z^{-3}} \quad (3.79) \]
Including the measurement noise, the model for the process is

\[ y(k) + 0.8825y(k - 1) = 0.47u(k - 2) + w(k) - 0.7w(k - 1) + 0.12w(k - 2); \quad (3.80) \]

where \( \{w\} \) is the white noise with \( N(0,0.2599) \) distribution.

The controller is a one-degree-of-freedom PID controller in the form

\[ K(z^{-1}) = \frac{k_0 + k_1 z^{-1} + k_2 z^{-2}}{1 - z^{-1}} \quad (3.81) \]

A set of well-tuned PID parameters is given as \( k_0 = 0.9102, k_1 = -0.9913, \) and \( k_2 = 0.1579. \)

The variance of the white noise is 0.2599, but is assumed to be unknown. Using the estimation algorithm based on a Laguerre network with time scale \( a = 0.3 \) and the number of stages \( N = 3 \), the estimated minimum variance \( \hat{\sigma}_{mv}^2 = 0.2724 \).

### 3.4.1 Case 1: Oscillation Caused by High Friction Valve

In this case, the loop controller is well-tuned, but the control valve exhibits high friction. The high friction of control valve is simulated by adding a nonlinear element into the control loop. The nonlinear element has the characteristics as

\[ v(t) = \begin{cases} 
0 & |\Delta u(t)| < \text{band} \\
\Delta u(t) & |\Delta u(t)| \geq \text{band}
\end{cases} \quad (3.82) \]

As a benchmark, the performance of the loop when the valve is normal is given. Figure 3.4 demonstrates the response (a) and spectrum (b) when the Controller and valve are in good condition.

When high friction occurs, the loop performance degrades and oscillation occurs. Table 3.2 shows the statistics when the valve is normal and the valve is of high friction. The
Loop Performance Monitoring

Figure 3.4: Response and spectrum when valve in normal condition

Performance index in the latter case indicates that the loop performance has degraded.

Table 3.2: Loop Performance Statistics For Case 1

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_{mv}^2$</th>
<th>$\sigma_y^2$</th>
<th>$\mu_y^2$</th>
<th>$PI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.2724</td>
<td>0.3329</td>
<td>0</td>
<td>0.1817</td>
</tr>
<tr>
<td>HighFriction</td>
<td>0.2724</td>
<td>0.6796</td>
<td>0.0146</td>
<td>0.5992</td>
</tr>
</tbody>
</table>

Having found the performance deterioration, the loop monitoring procedure then goes to check whether the loop is oscillating. Using spectral analysis, it detects that spectrum spikes exceeding a prescribed threshold, and therefore the occurrence of oscillation. The normalized oscillation frequencies are $f_1 = 0.0078$, $f_2 = 0.0273$, and $f_3 = 0.0625$. The Figure 3.5 shows the loop output (a) and its spectrum (b).

Upon detecting the oscillation, the procedure first checks the valve since a control valve is usually the trouble-maker. The RLS algorithm is used to estimate the valve input-output relationship and compare it with the one obtained when the valve was normal. Figure 3.6 illustrates the normal valve input-output relationship (the solid line)
Figure 3.5: Response and spectrum when valve in normal condition and the high friction valve input-output relationship (the dashed line).

Checking the estimated valve input-output relationship curve, the loop monitoring procedure finds that curve lies far from the acceptable region, then it makes the decision that the valve is in the state of high friction which causes the oscillation.

The above results given by the loop monitoring procedure show that the degradation of loop performance caused by the high friction of valve is successfully detected and diagnosed.

3.4.2 Case 2: Oscillation Caused by Poorly Tuned Controller

In this case, the control valve is in the normal condition but the PID controller is de-tuned and causes the loop to oscillate. The PID controller parameters, for both well-tuned and poorly-tuned cases, are listed in the Table 3.3

When the poorly-tuned controller is used, the degradation of loop performance is detected by the loop monitoring procedure. Refer to Table 3.4 for the related statistics.

After the performance degradation is detected, a check for loop oscillation occurs. It is
Figure 3.6: Response and spectrum when valve in high friction

Table 3.3: Detuned PID Controller Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>$k_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-tuned</td>
<td>0.9102</td>
<td>-0.9913</td>
<td>0.1579</td>
</tr>
<tr>
<td>Poorly-tuned</td>
<td>0.3987</td>
<td>-0.3654</td>
<td>0.1565</td>
</tr>
</tbody>
</table>

Table 3.4: Loop Performance Statistics For Case 2

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\sigma}_{mv}^2$</th>
<th>$\sigma_{y}^2$</th>
<th>$\mu_y^2$</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well - tuned</td>
<td>0.2724</td>
<td>0.3329</td>
<td>0</td>
<td>0.1817</td>
</tr>
<tr>
<td>Poorly - tuned</td>
<td>0.2724</td>
<td>0.9222</td>
<td>0</td>
<td>0.7046</td>
</tr>
</tbody>
</table>
found that there exist spectral spikes exceeding the preset threshold. Refer to Figure 3.7 for the response and spectrum of the loop output. The normalized oscillation frequency is 0.0391.

![Response and spectrum for poorly-tuned controller](image)

**Figure 3.7:** Response and spectrum for poorly-tuned controller

Next, the loop monitoring procedure checks whether the oscillation is caused by the valve. It estimates the valve input-output relationship curve and compare it with the reference. The estimated input-output relationship, as shown in Figure 3.8, falls within the predefined "normal region". The valve is determined to be in the normal condition. At this time, the loop monitoring procedure goes to check the controller by calculating the oscillation index, which is computed as $I_{osc} = 0.0682$. This means that the oscillation is indeed caused by the controller.

The results indicate that the loop monitoring procedure has successfully detected the performance degradation and diagnosed the factor which causes the degradation. In this case, the performance degradation is caused by the controller.
3.5 Summary

To assess the loop performance, the loop performance monitoring procedure described above uses the normalized performance index proposed by Desborough and Harris. To estimate the minimum possible variance for the performance index, a Laguerre network is used for the estimation. The procedure also uses spectral analysis to detect loop oscillation. To diagnose the causes for oscillation, an ANM-based valve non-linearity check method is also implemented to detect the presence of valve high friction. An oscillation index can be used to locate oscillation caused by a poorly or de-tuned controller. The simulation results from the two cases have shown that the loop performance monitoring procedure has the ability to detect and diagnose two major causes for the degradation of loop performance.
Chapter 4

Controller Tuning

4.1 Introduction

Poor process performance caused by the controller arises from two main causes. First, the controller is not well-tuned initially and so gives poor performance. Second, the dynamics of a process changes over time and causes an even well-tuned controller to become detuned and then lead to the degradation of performance. When the performance monitoring mechanism determines the controller is causing the poor loop performance, controller tuning is necessary so that the performance can be improved as soon as possible. Tuning a controller is not an easy task even for experienced process personnel. Ideally, controller tuning is based on the full knowledge of process dynamics, knowledge that is seldom available. Although some empirical methods, such as the Ziegler-Nichols frequency response methods, can be used to tune PID controllers, these methods are not always reliable. The side effect of oscillation introduced by these methods also has severe negative effects on the production. The iterative feedback tuning (IFT) method proposed by Hjalmarsson et al. [17, 21] has been tested and found to be suitable for loop controller tuning if the system is linear time-invariant. The method does not require a process model, and no assumptions on the process other than linearity and time-invariance are needed. For a slow process, given that the control valve is in normal condition, the above two conditions can usually be satisfied.
4.2 Iterative Feedback Tuning

Consider a single-loop unknown system illustrated in Figure 4.9

![Block diagram of the closed-loop system](image)

Figure 4.9: Block diagram of the closed-loop system

The process is described as

\[ y_t = Pu_t + v_t \quad (4.83) \]

where \( P \) is a linear time-invariant process transfer function, \( \{v_t\} \) is an unmeasurable zero mean weakly stationary random process, and \( t \) represents the discrete time interval.

Assume that this system is controlled by a two-degrees-of-freedom controller as

\[ u_t = C_r(\rho)r_t - C_y(\rho)y_t \quad (4.84) \]

where \( C_r(\rho) \) and \( C_y(\rho) \) are linear time-invariant transfer functions parameterized by some parameter vector \( \rho \in R^{n_p} \), and \( \{r_t\} \) is an external deterministic reference signal,
Controller Tuning

independent of \{v_t\}.

For the sake of convenience, the time argument of the signals will be omitted whenever signals are collected from the closed-loop system with the controller \(C(p) \triangleq \{C_r(p), C_y(p)\}\).

Let \(T_d\) be the reference model for closed-loop response from the reference signal to the output signal

\[ y_d = T_d r \] (4.85)

The error between the achieved and desired response is

\[ \tilde{y}(p) = y(p) - y_d = \frac{C_r(p)P}{1 + C_y(p)P} r - T_d r + \frac{1}{1 + C_y(p)P} v \] (4.86)

The error consists of a contribution due to tracking error of the reference signal \(r\) and an error due to the random disturbance.

For a controller with a certain fixed structure parameterized by \(p\), the control design objective can be formulated as a minimization of some norm of \(\tilde{y}\) over the controller parameter vector \(p\). The optimal controller parameter vector is defined by

\[ \rho^* = \text{arg}(\min J(p)) \] (4.87)

where \(J(p)\) is restricted to a quadratic criterion with the form

\[ J(p) = \frac{1}{2N} E\left[ \sum_{t=1}^{N} (L_y \tilde{y}(p))^2 + \lambda \sum_{t=1}^{N} (L_u u_t(p))^2 \right] \] (4.88)

where \(E[\cdot]\) denotes the expectation with respect to the weakly stationary disturbance \(v\). \(\lambda\) is the control penalty. \(L_y\) and \(L_u\) are the frequency weighting filters for the error between the desired response and the achieved response, and the control effort respectively.
The frequency weighting filters can be used to focus the attention of the controller on specific frequency bands in the input and/or output response of the closed loop system, for example, to suppress undesirable oscillations in these signals. Conversely, they can be used as notch filters in the frequency bands where the output is dominated by measurement noise.

To simplify the notation here, assume that $L_y = L_u = 1$. Then the optimal control parameter vector $v$ can be obtained by solving the following equation

$$
0 = \frac{\partial J}{\partial \rho} = \frac{1}{N} E[\sum_{t=1}^{N} \tilde{y}(\rho) \frac{\partial \tilde{y}}{\partial \rho}(\rho) + \lambda \sum_{t=1}^{N} u_t(\rho) \frac{\partial u_t}{\partial \rho}(\rho)]
$$

(4.89)

According to Hjalmarsson [17, 21], the following stochastic approximation iterative algorithm can be used to obtain the solution

$$
\rho_{i+1} = \rho_i - \gamma_i R_i^{-1} \text{est\left[}\frac{\partial J}{\partial \rho}(\rho_i)\text{\right]}
$$

(4.90)

where $R_i$ is some appropriate positive definite matrix, $\gamma_i$ is a variable positive real scalar that determines the step size, and $\text{est\left[}\frac{\partial J}{\partial \rho}(\rho_i)\text{\right]}$ is the estimate of the gradient.

In each iteration $i$ of the controller tuning algorithm, three experiments with each of duration $N$ must be applied to the process. Each experiment has the following corresponding reference signal

$$
r_i^1 = r; r_i^2 = r - y_i^1; r_i^3 = r.
$$

(4.91)

where $r_i^j$ is the $j$th experiment in the $i$th iteration, and $y_i^j$ is the corresponding output.

Hjalmarsson gives the estimate of the gradient $\frac{\partial J}{\partial \rho}(\rho_i)$ as

$$
\text{est\left[}\frac{\partial J}{\partial \rho}(\rho_i)\text{\right]} = \frac{1}{N} \sum_{t=1}^{N} \tilde{y}(\rho) \text{est\left[}\frac{\partial y_t}{\partial \rho}(\rho_i)\text{\right]} + \lambda u_t(\rho_i) \text{est\left[}\frac{\partial u_t}{\partial \rho}(\rho_i)\text{\right]}
$$

(4.92)

The partial derivatives $\frac{\partial y}{\partial \rho}(\rho_i)$ and $\frac{\partial u}{\partial \rho}(\rho_i)$ are computed by the following equations
Controller Tuning

\[ est[\frac{\partial y}{\partial \rho}(\rho_i)] = \frac{1}{C_r(\rho_i)} \left[ \frac{\partial C_r}{\partial \rho}(\rho_i) - \frac{\partial C_y}{\partial \rho}(\rho_i)y^3(\rho_i) + \frac{\partial C_y}{\partial \rho}y^2(\rho_i) \right] \]

and

\[ est[\frac{\partial u}{\partial \rho}(\rho_i)] = \frac{1}{C_r(\rho_i)} \left[ \frac{\partial C_r}{\partial \rho}(\rho_i) - \frac{\partial C_y}{\partial \rho}(\rho_i)u^3(\rho_i) + \frac{\partial C_y}{\partial \rho}u^2(\rho_i) \right] \]

It is important to note here that the usually unknown gradient can be obtained entirely from the input-output data collected from the actual closed loop system, by performing the experiments on the system.

The motivation of the third experiment is to make the estimate of the gradient \( \frac{\partial J}{\partial \rho} \) unbiased, i.e.,

\[ E\{est[\frac{\partial J}{\partial \rho}(\rho_i)]\} = \frac{\partial J}{\partial \rho}(\rho_i) \]

There are many possible choices for the matrix \( R_i \). The identity matrix is one choice, which gives the negative gradient direction. According to Hjalmarsson, however, the following \( R_i \) is considered to be a better choice from the standpoint of practical application.

\[ R_i = \frac{1}{N} \sum_{t=1}^{N} \left( est[\frac{\partial y_t}{\partial \rho}(\rho_i)]est[\frac{\partial y_t}{\partial \rho}(\rho_i)]^T + \lambda est[\frac{\partial u_t}{\partial \rho}(\rho_i)]est[\frac{\partial u_t}{\partial \rho}(\rho_i)]^T \right) \]

To guarantee the convergence of the iterative algorithm, the elements of the sequence \( \gamma_i \) must satisfy the following conditions:

1. \( \gamma_i \geq 0 \) and \( \sum_{i=1}^{\infty} \gamma_i = \infty \);
2. \( \sum_{i=1}^{\infty} \gamma_i^2 < \infty \).

When the process is linear and time-invariant, the IFT algorithm can converge to a local minimum, given a linear restricted complexity controller. To guarantee the parameters converge to the correct local minimum, an additional condition is that the controller preserved closed-loop stability before its being tuned.
4.3 Simulation

This section demonstrates the effects of using the above IFT algorithm to tune controllers, for the process as

\[ P(z^{-1}) = \frac{z^{-1} + 0.6z^{-2}}{1 - 1.8z^{-1} + 0.81z^{-2}} \]  \hspace{1cm} (4.97)

and the measurement noise \( \{v\} \) generated by

\[ v(t) = \omega_n(t) - 1.1\omega_n(t - 1) + 0.3\omega(t - 2) \]  \hspace{1cm} (4.98)

where \( \{\omega_n\} \) is the white noise with variance \( \sigma_v^2 = 0.2534 \).

In the following simulations, the weighting filter polynomials \( L_y \) and \( L_u \) are taken as \( L_y = L_u = 1 \), and the control penalty \( \lambda = 0 \).

4.3.1 Case 1

In this case, the controller is an one-degree-of-freedom PID controller with transfer function

\[ C(z^{-1}) = \frac{k_0 + k_1z^{-1} + k_2z^{-2}}{1 - z^{-1}} \]  \hspace{1cm} (4.99)

Since this is an one-degree-of-freedom controller, only the first two experiments are needed. For the first experiment, the reference signal is \( r = 200 \). For the second experiment, the reference signal is \( r - y^1 \), and \( \{y^1\} \) is the output sequence generated in the first experiment.

For regulatory control, refer to Table 4.5 for the PID parameters before tuning and after tuning, and refer to Figure 4.10 for the loop responses before tuning (a) and after tuning (b).
Table 4.5: One DOF PID Controller Parameters For Regulatory Control

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>$k_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before tuning</td>
<td>0.0350</td>
<td>-0.0400</td>
<td>0.0100</td>
</tr>
<tr>
<td>After tuning</td>
<td>0.3409</td>
<td>-0.5236</td>
<td>0.2030</td>
</tr>
</tbody>
</table>

Figure 4.10: Responses of regulatory control using one DOF PID controller before (a) and after (b) tuning
The controller parameters were fixed when the sixth iterations was complete. The output variances before and after the tuning is $\sigma^2_v = 5.0774$ and $\sigma^2_v = 0.3613$, respectively.

When the PID controller is used for setpoint tracking, the IFT algorithm also shows a good tuning result. Refer to Table 4.6 for the PID parameters before tuning and after tuning, and refer to Figure 4.11 for the loop responses before tuning (c) and after tuning (d).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Before tuning</th>
<th>After tuning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_0$</td>
<td>0.01</td>
<td>0.8465</td>
</tr>
<tr>
<td>$k_1$</td>
<td>-0.01</td>
<td>-1.1891</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.001</td>
<td>0.4984</td>
</tr>
</tbody>
</table>

Figure 4.11: Responses of setpoint tracking using one DOF PID controller before (a) and after (b) tuning.

The loop responses illustrated above show that the IFT algorithm has tuned the one-degree-of-freedom PID controller successfully.
### 4.3.2 Case 2

In this case, the controller is a two-degrees-of-freedom PID controller of the structure

\[(z - 1)(z - a)u(t) = (k_0 z^2 + k_1 z + k_2)r(t) - (s_0 z^2 + s_1 z + s_2)y(t)\]  \hspace{1cm} (4.100)

where \(a\) is the design parameter and is taken as \(a = 0.5\).

In this case, the controller has two degrees of freedom, therefore, all the three experiments are needed. For the first experiment, the reference signal is \(r = 200\). For the second experiment, the reference signal is \(r - y^1\), and for the third experiment, the reference signal is again \(r\). Note that \(\{y^1\}\) is the output sequence generated in the first experiment.

Refer to Table 4.7 for the PID parameters before tuning and after tuning for the regulatory control. Figure 4.12 illustrates the loop responses for regulatory control before tuning (a) and after tuning (b).

#### Table 4.7: Two DOF PID Controller Parameters For Regulatory Control

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(t_0)</th>
<th>(t_1)</th>
<th>(t_2)</th>
<th>(s_0)</th>
<th>(s_1)</th>
<th>(s_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before tuning</td>
<td>0.0030</td>
<td>-0.0030</td>
<td>0.0005</td>
<td>0.0030</td>
<td>-0.0030</td>
<td>0.0005</td>
</tr>
<tr>
<td>After tuning</td>
<td>1.3389</td>
<td>-2.3080</td>
<td>1.0615</td>
<td>0.5348</td>
<td>-0.8517</td>
<td>0.4092</td>
</tr>
</tbody>
</table>

The controller parameters were taken after ten iterations. For the regulatory control as illustrated in Figure 4.12, the output variances before and after the tuning are \(\sigma_y^2 = 4.4014\) and \(\sigma_y^2 = 0.9230\), respectively.

Table 4.8 lists the original and tuned parameters of the two-degrees-of-freedom PID controller when it is used for setpoint tracking. Figure 4.13 illustrates the loop responses for setpoint tracking control before tuning (a) and after tuning (b).
Figure 4.12: Responses of regulatory control using two DOF PID controller before (a) and after (b) tuning

Table 4.8: Two DOF PID Controller Parameters For Setpoint Tracking

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$t_0$</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$s_0$</th>
<th>$s_1$</th>
<th>$s_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before tuning</td>
<td>0.0030</td>
<td>-0.0030</td>
<td>0.0005</td>
<td>0.0030</td>
<td>-0.0030</td>
<td>0.0005</td>
</tr>
<tr>
<td>After tuning</td>
<td>1.3890</td>
<td>-2.8643</td>
<td>1.5631</td>
<td>0.5054</td>
<td>-0.8098</td>
<td>0.3922</td>
</tr>
</tbody>
</table>
Figure 4.13: Responses of setpoint tracking using two DOF PID controller before (a) and after (b) tuning

The responses illustrated above show that the performance of the two-degree-of-freedom PID controller has been improved greatly after being tuned by the IFT algorithm.

4.4 Summary

The IFT controller tuning algorithm possesses the advantage of not needing to know the process dynamics. The whole tuning process depends only on the data acquired from the experiments. Given a linear controller with restricted complexity structure, it will converge to a local minimum of the design criterion if the process is linear and time-invariant. This direct optimal tuning algorithm is particularly suitable for tuning basic control loops in process industries, which typically use PID controllers. Compared with the empirical Ziegler-Nichols frequency PID tuning methods, the IFT algorithm does
not introduce oscillation to the process. Although it requires more data, it offers a well-tuned PID controller which usually gives faster achieved response. The above simulations demonstrate an overall increase of performance for all cases by tuning the controllers with the IFT algorithm. Although the simulations show the "ringing" phenomenon in the responses after tuning, it can be eliminated by designing a notch filter as the frequency weighting filter $L_y$. 
5.1 Introduction

The use of PCA and PLS for performance monitoring has been described in Chapter 2. The PCA- or PLS-based performance monitoring techniques are very suitable for monitoring the process at the supervisory level. The loop performance monitoring scheme, as described in Chapter 3, is able to assess an individual loop performance and provide detailed diagnosis of some common causes for loop performance degradation, such as loop oscillation, valve high friction and bad controller tuning. Both performance monitoring schemes have drawbacks. PCA- and PLS-based monitoring techniques are poor at diagnosing process abnormality. The loop performance monitoring mechanism is essentially limited to a single loop. For a typical process with multiple control loops, neither of the techniques is sufficient to provide an effective performance monitoring and diagnosis approach. For process industries, however, the performance monitoring system should be able to provide process-wide performance monitoring and trouble-shooting. This requires the performance monitoring system to integrate multiple performance monitoring and fault diagnosis techniques systematically. Driven by the motivation of providing process-wide performance monitoring and fault diagnosis, this chapter proposes a hierarchically architectured performance monitoring mechanism, which combines the previously introduced PCA- and PLS-based monitoring techniques with the loop performance monitoring scheme. This system is intended to provide both overall process monitoring
to help process personnel to see the status of their process and to provide an effective diagnosis ability for some common causes of degradation of process performance.

5.2 Hierarchically Architectured Performance Monitoring System

The proposed performance monitoring system consists of two levels, a higher-level subsystem and a cluster of lower-level subsystems for each monitored loop. The higher-level subsystem aims at providing overall monitoring for a multi-loop process and a comprehensive diagnosis in the event of the occurrence of process abnormality. Information is extracted from the process operating data and the diagnosis coming from the lower-level monitoring subsystem. In this higher-level subsystem, the PCA- and PLS-based techniques are deployed to handle a large amount of correlated process data from different loops and to provide reliable fault detection. In addition, a comprehensive analysis mechanism is located at this level. The comprehensive analysis mechanism provides some diagnostic information to the lower-level subsystems when a fault is detected and also undertakes a comprehensive analysis based on the information extracted from operating data and the diagnosis results from the lower-level subsystems so that a comprehensive process diagnosis can be achieved. The purpose of a lower-level subsystem is to provide loop-oriented monitoring and diagnosis, based on the loop operating data and the diagnostic information from the high-level subsystem. Its main responsibilities are to assess loop performance, detect and locate loop oscillation, check valve high friction, and evaluate the controller.

The two level subsystems are combined in an integrated way and monitor the process and diagnose the causes responsible for the degradation of process performance. The overview of the proposed hierarchical process performance monitoring system is illustrated in Figure 5.14.
The PCA- and PLS-based monitoring procedures use the operating data from either the whole process or from multiple key process loops. The data will be analysed by either PCA or PLS to detect when process abnormality occurs. The decision is made based on the "normal operating boundary" defined by the statistics of the reference model built in advance. If abnormality is found in the process, the comprehensive analysis mechanism will send the diagnostic information to the concerned loops to indicate that an abnormal event is present. The monitoring procedure is illustrated in Figure 5.15.

When the high-level monitoring procedure detects some abnormal process event, it will send the detection results to the comprehensive diagnosis mechanism. The identification of event source is then started and the diagnosis command is sent to the loop monitoring procedures in the lower-level subsystems.

The loop performance monitoring procedure in the lower level will initiate the diagnosis procedure when the message from the higher-level subsystem is received or when...
Integrated Performance Monitoring Procedure

Figure 5.15: Procedure for higher-level performance monitoring

degradation of loop performance is detected. The diagnosis procedure is demonstrated in Figure 5.16

The diagnosis results will be reported to the comprehensive analysis mechanism for further processing. The upward messages from each loop performance monitoring procedure include wholly or partly the items listed in Table 5.9

Table 5.9: Upward Messages Format

<table>
<thead>
<tr>
<th>Items</th>
<th>Performance</th>
<th>Oscillation</th>
<th>Valve</th>
<th>Controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>0/1</td>
<td>0/1</td>
<td>0/1</td>
<td>0/1</td>
</tr>
</tbody>
</table>

In Table 5.9, the message "0" means the normal condition, such as good performance, non-existence of oscillation, valve normal or controller well-tuned; and the message "1" means abnormal condition, such as degradation of performance, existence of oscillation, high friction of valve or poor tuning of controller.
The comprehensive analysis mechanism does further analysis, based on the information given by PCA or PLS analysis and the information is sent back by the related loop performance monitoring schemes. Then, the results will be sent to the process personnel for corrective actions to be taken.

In essence, the comprehensive analysis mechanism acts as the interface between the high-level subsystem and the lower-level subsystem.

### 5.3 Simulation

In the following simulations, a three loop system, shown in Figure 5.17, is used as the testbed. This is actually a blend tank level control system, which consists of two inner flow control loops, one for hardwood and the other for softwood, and one outer level control loop. In the shown system, $y_1$ and $y_2$ are flows; $y_3$ is the tank level. $u_1$ and $u_2$ are signals to valve; $u_3$ is a flow setpoint. The controllers used to control the three loops are
discrete one-degree-of-freedom PID controllers. Three cases are simulated. In the first case, the control valve in loop 1 is given high friction which causes the loop to oscillate. In the second case, the controller in loop 3 is badly tuned and also causes an oscillation. In the last case, the controller in loop 1 is de-tuned and causes the degradation of loop performance.

![Figure 5.17: Block diagram of a blend tank level control system](image)

The transfer functions for the valves and the tank are given as below

Valve 1 in loop 1:

\[
P_1(s) = \frac{4}{1 + 8s}e^{-2s} \quad (5.101)
\]

Valve 2 in loop 2:

\[
P_2(s) = \frac{3}{1 + 7s}e^{-2s} \quad (5.102)
\]

Tank in loop 3:
\[ P3(s) = \frac{0.005}{s} \] (5.103)

In the above loops, \( v_i, i = 1, 2, 3 \) are the sequences of white noise with variance 0.36. The sampling interval for the three loops is 1 second. The discrete one-degree-of-freedom PID controller has the form as

\[ u(t) = u(t-1) + k_0 e(t) + k_1 e(t-1) + k_2 e(t-2) \] (5.104)

In the higher-level subsystem, the monitoring procedure based on PCA is used. A reference model, which defines the boundary of the normal operating status for the process, is built in advance from a set of data collected from the above system when it is in normal condition. The data are the measurements of the three loop outputs, which are illustrated in Figure 5.18, under normal condition.

![Figure 5.18: Normal responses of the three loops](image)

A segment of data with length \( n = 300 \) is used for the model building. Autoscaling
is used for the data pretreatment. Applying PCA to the reference data, the principal component number versus the explained variance is shown in Table 5.10

<table>
<thead>
<tr>
<th>Principal Component Number</th>
<th>Eigenvalue of Cov(X)</th>
<th>% Variance Captured by ThisPC</th>
<th>% Variance Captured Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.04e + 00</td>
<td>67.92</td>
<td>67.92</td>
</tr>
<tr>
<td>2</td>
<td>7.55e - 01</td>
<td>25.17</td>
<td>93.10</td>
</tr>
<tr>
<td>3</td>
<td>2.07e - 01</td>
<td>6.90</td>
<td>100</td>
</tr>
</tbody>
</table>

According to Table 5.10, two principal components can explain 93.1% of variance. Therefore, the number of principal components is chosen as two. With the confidence level $\alpha = 0.05$, the control limits for the normal operating region in terms of the Hotelling’s statistic and SPE statistic are obtained as $T^2_{2,285,0.05} = 6.0701$ and $SPE = 3.8415$.

The normal operating region can be represented graphically using Anderson’s elliptical contour. The plot of scores for another segment of the reference data is shown in Figure 5.19.

And the $SPE$ plot and $T_2$ plot are illustrated in Figure 5.20 and Figure 5.21, respectively.

The plots above indicate that the model is valid. This reference model will be used for the following simulations.

**5.3.1 Case 1: Oscillation caused by high friction valve**

In this case, the control valve in loop 1 is set to be in the state of high but unknown friction. The controller in loop 1 still uses the well-tuned parameters. The loop responses given by the three loops are illustrated in Figure 5.22.
Figure 5.19: Plot of scores for reference data

Figure 5.20: Plot of SPE for reference data
Figure 5.21: Plot of $T^2$ for reference data

Figure 5.22: Responses of the three loops for case 1
In the higher-level, a sliding window is used by the monitoring procedure to project the new data into the coordinate system defined by the model. The plot of new scores for a subset of data with length \( n = 150 \) is shown in Figure 5.23.

![Figure 5.23: Plot of new scores for case 1](image)

The corresponding \( T^2 \) and \( SPE \) statistics are demonstrated in Figure 5.24 and Figure 5.25, respectively.

The procedure compares the \( T^2 \) and \( SPE \) statistics with their corresponding control limits for each observation. Once it detects that either of the control limits is exceeded, it indicates alarm and sends downward messages to the loop monitoring procedures in the lower-level subsystem. The plots indicate that the occurrence of process abnormality has been detected by the monitoring procedure in the higher-level subsystem. The monitoring procedure monitors every new observation. At each time when the high-level procedure detects that a new observation is outside any of control limits, it gives an alarm signal. Meanwhile, a diagnosis command will be sent to the loop monitoring procedures in the
Figure 5.24: Plot of $T^2$ for case 1

Figure 5.25: Plot of SPE for case 1
loops concerned. For this case, all the three loop monitoring procedures will be sent the command.

At the same time, each loop monitoring procedure in the lower-level subsystem is also assessing the performance of its own loop. Table 5.11 lists the statistics for the three loops.

Table 5.11: Loop Statistics For Case 1

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_{mv}^2$</th>
<th>$\sigma_y^2$</th>
<th>$\mu_y^2$</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop1</td>
<td>0.3958</td>
<td>3.0663</td>
<td>0.0015</td>
<td>0.8710</td>
</tr>
<tr>
<td>Loop2</td>
<td>0.3765</td>
<td>0.4365</td>
<td>0.0066</td>
<td>0.1503</td>
</tr>
<tr>
<td>Loop3</td>
<td>0.3961</td>
<td>0.5264</td>
<td>0.01</td>
<td>0.2614</td>
</tr>
</tbody>
</table>

The loop monitoring procedures for loop 1 and loop 3 have detected the performance degradation. Then, they check whether oscillation exists in their own loops. The result of spectral analysis given by the procedure in loop 1, as shown in Figure 5.26, indicates that an oscillation exists in this loop. Figure 5.27 shows the spectrum of loop 3 output, which indicates that no oscillation exists in loop 3.

Based on the result of spectral analysis, the next step in loop 1 is to check the valve by estimation the input-output relationship curve of the valve. Figure 5.28 illustrates the reference and the actual valve input-output relationship.

The estimated actual valve input-output relationship curve shows that high friction has occurred to the valve. This implies that the oscillation is generated by the valve within this loop.

After obtaining the diagnosis result, the monitoring procedure in loop 1 sends the upward messages as listed in Table 5.12 to the comprehensive analysis mechanism in the high-level monitoring subsystem.

For the procedure in loop 3, it sends the upward messages as listed in Table 5.13
Figure 5.26: Spectrum for loop 1: case 1

Figure 5.27: Spectrum for loop 3: case 1
Figure 5.28: Valve input-output relationship: estimated (dash line) vs. reference (solid line): case 1

Table 5.12: Upward Messages From Loop 1 For Case 1

<table>
<thead>
<tr>
<th>Items</th>
<th>LoopNo</th>
<th>Performance</th>
<th>Oscillation</th>
<th>Valve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Since no performance degradation has been found in loop 2, its monitoring procedure sends the upward messages as listed in Table 5.14 to the higher-level subsystem to inform that no gradation of performance has occurred to its loop.

Table 5.14: Upward Messages From Loop 2 For Case 1

<table>
<thead>
<tr>
<th>Items</th>
<th>LoopNo.</th>
<th>Performance</th>
<th>Oscillation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

When the comprehensive diagnostic mechanism receives all the reports from the three loop monitoring procedures, it concludes that the first trouble maker should be the valve in loop 1. Then, it sends this result to the process personnel.

5.3.2 Case 2: Performance degradation caused by poorly-tuned controller

In this case, the valves in both loop 1 and loop 2 are set to be normal and the two loop PID controllers are well-tuned. The controller in loop 3, however, is poorly tuned to cause oscillation. The well tuned and poorly tuned parameters of loop 3 PID controller are listed in Table 5.15.

The responses of the three loops in this situation are shown in Figure 5.29.

In the higher-level, the PCA-based monitoring procedure also uses a sliding window to project the new data into the new coordinate system to monitor the process. The plot of new scores for a subset of data with length $n = 150$ is shown in Figure 5.30.
Table 5.15: Well and Poorly Tuned Parameters of Loop 3 PID Controller

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>$k_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well Tuned</td>
<td>0.9135</td>
<td>-0.7231</td>
<td>0.1012</td>
</tr>
<tr>
<td>Poorly Tuned</td>
<td>3.6750</td>
<td>-1.1410</td>
<td>0.2379</td>
</tr>
</tbody>
</table>

Figure 5.29: Responses of the three loops for case 2
The corresponding $T^2$ and $SPE$ statistics are demonstrated in Figure 5.31 and Figure 5.32, respectively.

The calculated $T^2$ and $SPE$ statistics for each observation are compared with their corresponding control limits. The plots indicate that the occurrence of process abnormality has been detected by the monitoring procedure in the higher-level subsystem. When process abnormality is detected, the higher-level monitoring procedure alarms the process personnel and sends diagnosis command to each of the loop monitoring procedures in the lower-level subsystem.

Meanwhile, each of the loop monitoring procedures in the lower-level subsystem is also assessing the performance of its own loop. Table 5.16 lists the assessment results of the three loops.

All the loop monitoring procedures in the three loops have detected the degradation of performance. They then turn to check whether oscillations exist in their loops. The
Figure 5.31: Plot of $T^2$ for case 2

Figure 5.32: Plot of SPE for case 2
Table 5.16: Loop Statistics For Case 2

<table>
<thead>
<tr>
<th>Loop</th>
<th>$\sigma^2_{mv}$</th>
<th>$\sigma^2_y$</th>
<th>$\mu^2_y$</th>
<th>PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop1</td>
<td>0.3986</td>
<td>0.6786</td>
<td>0.0006</td>
<td>0.4131</td>
</tr>
<tr>
<td>Loop2</td>
<td>0.3728</td>
<td>0.5684</td>
<td>0.0033</td>
<td>0.3478</td>
</tr>
<tr>
<td>Loop3</td>
<td>0.4483</td>
<td>0.6715</td>
<td>0.0008</td>
<td>0.3324</td>
</tr>
</tbody>
</table>

results of spectral analysis for the three loop are shown in Figure 5.33, Figure 5.34, and Figure 5.35.

The results of spectral analysis, given by the three loop monitoring procedures, indicate that oscillations exist in all of the three loops. The normalized frequencies for the oscillations all the three loops are found to be 0.0273.

After having detected the oscillation, the monitoring procedure in each loop then begins to locate whether the detected oscillation is generated by its own loop. Since loop1 and loop 2 have valves involved, their monitoring procedures first evaluate the
Figure 5.34: Spectrum for loop 2: case 2

Figure 5.35: Spectrum for loop 3: case 2
Integrated Performance Monitoring Procedure

valves. Figure 5.36 and Figure 5.37 illustrate the results of valve evaluation for loop 1 and loop 2, respectively.

![Figure 5.36: Valve input-output relationship: estimated (dash line) vs. reference (solid line): case 2](image)

The estimated input-output relationship curves for both loop valves fall within the prescribed normal region. Therefore, both the loop monitoring procedures reach the conclusion that the valve is normal. Just as the monitoring procedure in loop 3, the procedures in loop 1 and loop 2 then go to check their controller. Table 5.17 shows the oscillation indices given by each loop monitoring procedure.

<table>
<thead>
<tr>
<th>Loop</th>
<th>Loop1</th>
<th>Loop2</th>
<th>Loop3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>0.0273</td>
<td>0.0273</td>
<td>0.0273</td>
</tr>
<tr>
<td>$I_{osc}$</td>
<td>0.5212</td>
<td>0.5183</td>
<td>0.0068</td>
</tr>
</tbody>
</table>
The calculated oscillation indices show that only loop 3 can be generating the oscillation within its loop by the controller. After obtaining the diagnosis results, the procedures then send upward messages to the high-level subsystem. The upward messages from each loop are as listed in Table 5.18, Table 5.19 and Table 5.20.

<table>
<thead>
<tr>
<th>Items</th>
<th>LoopNo.</th>
<th>Performance</th>
<th>Oscillation</th>
<th>Valve</th>
<th>Controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

When all the upward messages are received by the comprehensive diagnostic mechanism, it identifies the possible source for the abnormal event as the controller in the loop 3. Then, it informs the process personnel of the diagnostic result and commands the monitoring procedure in loop 3 to undertake the controller tuning.
Table 5.19: Upward Messages From Loop 2 For Case 2

<table>
<thead>
<tr>
<th>Items</th>
<th>LoopNo</th>
<th>Performance</th>
<th>Oscillation</th>
<th>Valve</th>
<th>Controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.20: Upward Messages From Loop 3 For Case 2

<table>
<thead>
<tr>
<th>Items</th>
<th>LoopNo</th>
<th>Performance</th>
<th>Oscillation</th>
<th>Controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

5.3.3 Case 3: Performance degradation caused by de-tuned controller

In this case, the valves in both loop 1 and loop 2 are set to be normal and the PID controllers in loop 2 and loop 3 are well-tuned, but the controller in loop 1 is de-tuned and leads to the degradation of performance. The original and de-tuned parameters of loop 1 PID controller are listed in Table 5.21.

Table 5.21: Original and De-tuned Parameters of Loop 1 PID Controller

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>$k_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.4822</td>
<td>-0.5948</td>
<td>0.1121</td>
</tr>
<tr>
<td>De-tuned</td>
<td>0.0287</td>
<td>-0.0254</td>
<td>0.0228</td>
</tr>
</tbody>
</table>

The responses of the three loops under this condition are shown in Figure 5.38.

The higher-level monitoring subsystem has detected the abnormality of process based on the new score distance and SPEs. Then it starts alarming and sends diagnosis command to each of the loop monitoring procedures in the lower-level. The plot of scores for a subset of data with length $n = 150$ is shown in Figure 5.39.

The corresponding $SPE$ and $T^2$ statistics are demonstrated in Figure 5.40 and Figure 5.41, respectively.
Figure 5.38: Responses of the three loops for case 3

Figure 5.39: Plot of new scores for case 3
Figure 5.40: Plot of SPE for case 3

Figure 5.41: Plot of $T^2$ for case 3
The loop monitoring procedures in the lower-level subsystem are also assessing the performance. Table 5.16 lists the assessment results of the three loops.

Table 5.22: Loop Statistics For Case 3

<table>
<thead>
<tr>
<th></th>
<th>( \sigma_{xy}^2 )</th>
<th>( \sigma_y^2 )</th>
<th>( \mu_y^2 )</th>
<th>( PI )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop1</td>
<td>0.3978</td>
<td>0.5136</td>
<td>0.0004</td>
<td>0.2260</td>
</tr>
<tr>
<td>Loop2</td>
<td>0.3982</td>
<td>0.4542</td>
<td>0.0006</td>
<td>0.1245</td>
</tr>
<tr>
<td>Loop3</td>
<td>0.4083</td>
<td>0.5453</td>
<td>0.0003</td>
<td>0.2512</td>
</tr>
</tbody>
</table>

While the loop 2 monitoring procedure has found the performance to be normal, the monitoring procedures in loop 1 and loop 3 have indeed detected the degradation of performance. Then, they do spectral analysis to check whether their loop are oscillating. Figure 5.42 and Figure 5.43 show the spectrum of the two loop outputs. The results of spectral analysis indicate that no oscillations exist in the two loops.
At this moment, all the loop monitoring procedures send upward messages to the higher-level subsystem to report their diagnosis. The loop 1 and loop 3 procedures also request to undertake controller tuning. The upward messages are as listed in Table 5.23, Table 5.24 and Table 5.25.

Table 5.23: Upward Messages From Loop 1 For Case 3

<table>
<thead>
<tr>
<th>Items</th>
<th>LoopNo.</th>
<th>Performance</th>
<th>Oscillation</th>
<th>TuningRequest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messages</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Based on the upward messages, the comprehensive diagnostic mechanism in the higher-level subsystem then first sends controller tuning command to the loop 1 monitoring procedure. Then, the loop 1 monitoring procedure starts to tune the PID controller using the embedded IFT tuning algorithm. The parameters of loop 1 PID controller before tuning and after tuning are listed in Table 5.26.
When the tuning process stops, the parameters of loop 1 PID controller are updated. The monitoring system continues to monitor the process performance. The new process responses are shown in Figure 5.44.

The plots of scores, $T^2$ and $SPE$ are illustrated in Figure 5.45, Figure 5.46, and Figure 5.47.

After tuning, the higher-level monitoring subsystem detects no process abnormality. The process statistics given by the lower-level subsystems are listed in Table 5.27. All the loop monitoring procedures obtain a positive results of performance assessment.

The simulation results have shown that the degradation of performance caused by the de-tuned loop 1 controller has been successfully detected by the monitoring system. The embedded IFT tuning algorithm has also successfully tuned the loop controller.

### Table 5.26: Parameters of Loop 1 PID Controller Before and After Tuning

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>$k_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before tuning</td>
<td>0.0287</td>
<td>-0.0254</td>
<td>0.0228</td>
</tr>
<tr>
<td>After tuning</td>
<td>0.4712</td>
<td>-0.5677</td>
<td>0.1014</td>
</tr>
</tbody>
</table>
Figure 5.44: Responses after tuning for case 3

Figure 5.45: Plot of new scores for case 3
Figure 5.46: Plot of new $SPE$ for case 3

Figure 5.47: Plot of new $T^2$ for case 3
Table 5.27: New Loop Statistics For Case 3

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\sigma}_{mx}^2$</th>
<th>$\sigma_y^2$</th>
<th>$\mu_y^2$</th>
<th>$PI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop1</td>
<td>0.3676</td>
<td>0.4495</td>
<td>0.0045</td>
<td>0.1904</td>
</tr>
<tr>
<td>Loop2</td>
<td>0.3601</td>
<td>0.4278</td>
<td>0.0001</td>
<td>0.1584</td>
</tr>
<tr>
<td>Loop3</td>
<td>0.4072</td>
<td>0.4832</td>
<td>0.0000</td>
<td>0.1573</td>
</tr>
</tbody>
</table>

5.4 Summary

Organizing a system hierarchically is a traditional and effective way to handle a complex problem. The hierarchical performance monitoring system proposed here combines the monitoring techniques described in the previous chapters in an integrated way so that they can monitor the process jointly and complementarily. The monitoring system has two advantages. First, it can provide overall process monitoring. Second, it is not only able to detect process abnormality reliably and diagnose the original fault causes, but can also provide mill-wide performance monitoring. The simulations have shown good performance by the monitoring system. Although the three cases for the simulation showed situations where both the high-level and lower-level monitoring procedures have detected the degradation of performance, there also exist situations in which lower-level monitoring procedures fail to detect the degradation of performance that is detected by the higher-level monitoring procedure alone. In such case, the monitoring system configured in a hierarchical manner is superior to single-mode monitoring systems.
Chapter 6

Conclusion and Future Work

6.1 Conclusion

Although distributed control systems (DCS) are familiar to most process engineers, constructing a performance monitoring system in a hierarchically distributed way is still not common. The performance monitoring system presented in this thesis is constructed with two hierarchical levels in an integrated way. The higher-level subsystem and the loop-oriented lower-level subsystem work jointly to fulfill the tasks of process monitoring and diagnosis. This performance monitoring system has the following features and advantages:

1. It consists of two hierarchical levels and provides process-wide performance monitoring. It can provide both overall performance monitoring and detailed diagnosis of process events.

2. The higher-level subsystem is PCA and PLS-based and capable of handling a large number of correlated process data set. This capacity makes the system capable of monitoring complex process plants.

3. The lower-level subsystem is loop-oriented. It is able to provide good performance assessment and detailed process diagnosis, such as loop oscillation detection and location, high friction of valves detection, and loop controller evaluation.

4. A controller tuning algorithm is embedded into the loop monitoring procedure, which can timely adjust a poorly tuned or de-tuned controller without knowledge of the
process dynamics.

5. Real-time performance monitoring can be easily implemented based on this system architecture.

6. This monitoring system can be configured flexibly in accordance with the process scale.

Compared with the commonly used performance monitoring techniques, this hierarchically constructed system is more suitable for comprehensive process monitoring.

6.2 Future Work

Since process performance monitoring is a relatively new field of research and is under development, there is as yet no technique that can be universally used with satisfactory performance. Although the performance monitoring mechanism presented in this thesis proposed a way to monitor a multi-loop process, further refinement is still needed to have it work more effectively. Some aspects for further work are addressed below:

1. A more robust PCA or PLS algorithm is needed to achieve more reliable analysis for data which contain colored noise.

2. The process-oriented reasoning mechanism in the higher-level subsystem needs to be refined. The diagnosis information yielded by the reasoning mechanism should give a comprehensive analysis for a process with many loops and variables, based on the information given by the PCA/PLS and the loop monitoring procedures in the lower level. To achieve this, an expert system approach can be considered.

3. For some processes, the time delay can be time varying. The estimation of time-varying time delay is important for an accurate performance benchmark so that a reliable loop performance assessment can be achieved. Such an estimation algorithm needs to be investigated.
Conclusion and Future Work

4. A more practical benchmark is needed for controller evaluation. The benchmark should take the control constraints into consideration. An LQG benchmark is a promising choice.

5. Model-free controller tuning algorithms, which can be used for non-linear controllers such as adaptive controllers or linear controllers with loose constraint on the linearity of systems, are worthy of further investigation.
Bibliography


87


