OPTIMAL CLOSED-LOOP EXPERIMENTS FOR ACCURATE STEP RESPONSE MODEL IDENTIFICATION

by

Warren Scott Ferguson

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Graduate Department of Chemical Engineering & Applied Chemistry
University of Toronto
©Copyright by Warren Scott Ferguson(2001)
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-62913-9
Abstract

OPTIMAL CLOSED-LOOP EXPERIMENTS FOR ACCURATE STEP RESPONSE MODEL IDENTIFICATION

Master of Applied Science, 2001

Warren Scott Ferguson

Department of Chemical Engineering and Applied Chemistry

University of Toronto

This thesis presents a new approach to the design of optimal closed-loop identification experiments. The proposed approach utilizes periodic excitation in a two-phase experiment to generate accurate step response estimates of the process. The two-phase approach only requires an a priori estimate of the process settling time, a significant improvement over designs presented in the literature which require knowledge of the process dynamics.

With a carefully designed test signal, preliminary information required for optimal identification of the process step response is collected in the first phase of the experiment. It is shown that periodic excitation coupled with the Frequency Sampling Filter (FSF) model structure allows for a highly efficient set of computations that can be performed on-line. The second phase of the experiment is used to ensure that the data collected from the complete experiment satisfies the optimality condition.
In memory of my Grandmother, Thelma V. Davies
(August 21, 1914 - July 27, 2001)
Acknowledgements

There are many people I would like to thank for contributing to my progress through four years of undergraduate and three years of graduate studies. Thank you friends, family, faculty, and in particular:

- Professor Will Cluett for his keen interest, support, and insightful discussions throughout my career as a student.

- Past and present members of the process control group, some of whom I know better than others, but to all of whom I owe gratitude for their companionship and support.

- Professor John Shaw for introducing the idea of graduate studies to me, a choice that has proven to be very rewarding.

- Natural Sciences and Engineering Research Council of Canada for their financial support.

- My family: mother, father, and grandmother. To say that they were supportive and encouraging throughout my education would be an understatement. I dedicate this thesis to them.

- Carla Freire, a special friend who celebrated my progress with me, and endured the hardships. The end of one chapter marks the beginning of another.
Contents

1 Introduction .......................................................... 1
  1.1 Motivation ..................................................... 1
  1.2 Preliminaries ................................................ 3
  1.3 Literature Review ........................................... 4
    1.3.1 Closed-Loop System Identification .................. 4
    1.3.2 Experimental Design .................................. 7
  1.4 Outline of the Thesis ....................................... 9

2 Frequency Sampling Filter Model Identification ........... 11
  2.1 Frequency Sampling Filter (FSF) Model Derivation .... 11
  2.2 Formulation of the Least Squares Problem ............... 13
  2.3 Step Response Models ...................................... 15
  2.4 Confidence Bounds ......................................... 16
    2.4.1 FSF Model Confidence Bounds ....................... 17
    2.4.2 FSF-generated Step Response Model Confidence Bounds . 17
  2.5 Optimal FSF Input Spectrum ................................ 18
  2.6 Multisine Design for FSF Model Identification .......... 20
  2.7 Summary .................................................... 22

3 The Role of the Noise Model .................................. 23
  3.1 Bias and Variance Effects .................................. 23
    3.1.1 Bias Effects .......................................... 23
    3.1.2 Variance Effects ................................... 25
  3.2 Simulation Studies .......................................... 26
    3.2.1 Simultaneous Estimation of Process and Noise Models . 27
    3.2.2 Prefiltering with an Incorrect Noise Model ........ 31
    3.2.3 Prefiltering with the True Noise Model ............ 32
    3.2.4 Quantifying the Results ............................. 33
  3.3 The Use of Averaging to Estimate the Noise Model ....... 35
  3.4 Summary ................................................... 42
## 4 The Two-Phase Approach to Optimal Closed-Loop Experiments

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Optimal Test Signal Design</td>
<td>43</td>
</tr>
<tr>
<td>4.2</td>
<td>Phase One</td>
<td>47</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Estimation of the Weighting Function</td>
<td>47</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Estimation of the Noise Model</td>
<td>49</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Estimation of the Background Spectrum</td>
<td>49</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Design of Phase 1 Excitation</td>
<td>51</td>
</tr>
<tr>
<td>4.3</td>
<td>Phase Two</td>
<td>54</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Fixing the Total Input Energy</td>
<td>55</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Design of Phase 2 Excitation</td>
<td>57</td>
</tr>
<tr>
<td>4.4</td>
<td>Simulation Studies</td>
<td>59</td>
</tr>
<tr>
<td>4.4.1</td>
<td>A Step-by-Step Example</td>
<td>59</td>
</tr>
<tr>
<td>4.4.2</td>
<td>A Comparative Example</td>
<td>68</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Monte Carlo Simulations</td>
<td>71</td>
</tr>
<tr>
<td>4.5</td>
<td>Possibilities and Limitations</td>
<td>75</td>
</tr>
<tr>
<td>4.5.1</td>
<td>The Transition Point</td>
<td>75</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Finite Length Data Sets</td>
<td>78</td>
</tr>
<tr>
<td>4.5.3</td>
<td>Non-Stationary Noise</td>
<td>82</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>85</td>
</tr>
</tbody>
</table>

## 5 Conclusions and Recommendations

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Conclusions</td>
<td>88</td>
</tr>
<tr>
<td>5.2</td>
<td>Recommendations</td>
<td>89</td>
</tr>
</tbody>
</table>

A Linking the Correlation Matrix to the Periodogram

B Transformations of FSF Spectra

C Magnitude of the Cross-Spectrum
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Optimal FSF Spectrum Ratios</td>
<td>19</td>
</tr>
<tr>
<td>3.1</td>
<td>Calculated ( \Gamma ) Values</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Average Estimated Noise Spectrum: Stationary Noise Case</td>
<td>38</td>
</tr>
<tr>
<td>3.3</td>
<td>Average Estimated Noise Spectrum: Non-Stationary Noise Case</td>
<td>40</td>
</tr>
<tr>
<td>4.1</td>
<td>Phase One Estimate of the Weighting Function</td>
<td>61</td>
</tr>
<tr>
<td>4.2</td>
<td>Phase One Estimate of the Noise Model</td>
<td>62</td>
</tr>
<tr>
<td>4.3</td>
<td>Phase One Estimate of the Background Spectrum</td>
<td>63</td>
</tr>
<tr>
<td>4.4</td>
<td>Phase One Input Energy</td>
<td>63</td>
</tr>
<tr>
<td>4.5</td>
<td>Specification of the Desired FSF Spectrum</td>
<td>64</td>
</tr>
<tr>
<td>4.6</td>
<td>Desired Phase Two Reference Spectrum</td>
<td>64</td>
</tr>
<tr>
<td>4.7</td>
<td>Comparison of the Desired and Actual FSF Spectra</td>
<td>68</td>
</tr>
<tr>
<td>4.8</td>
<td>FSF Spectrum Energies</td>
<td>70</td>
</tr>
<tr>
<td>4.9</td>
<td>Optimal and Achieved ( SE ) Values</td>
<td>70</td>
</tr>
<tr>
<td>4.10</td>
<td>Phase One GLS Estimate of the Weighting Function</td>
<td>72</td>
</tr>
<tr>
<td>4.11</td>
<td>FSF Spectrum Energies Based on GLS Estimates of the Weighting Function</td>
<td>72</td>
</tr>
<tr>
<td>4.12</td>
<td>Phase One SLS Estimate of the Weighting Function</td>
<td>73</td>
</tr>
<tr>
<td>4.13</td>
<td>FSF Spectrum Energies Based on SLS Estimates of the Weighting Function</td>
<td>75</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Schematic Diagram of a Closed-loop System .................. 4
2.1 Schematic Diagram of the FSF Model .......................... 12
3.1 Reference Signal Used in Simulation Studies ................... 27
3.2 Input-Output Data Collected Under Pure Feedback ............... 28
3.3 Step Response Estimates Using No External Excitation, estimates (-), true response (- -), 99% confidence bounds (-) ............... 29
3.4 Average Step Response Estimate Using No External Excitation, estimate (-), true response (- -), 99% confidence bounds (-) .......... 29
3.5 Input-Output Data In The Presence of External Excitation ........ 30
3.6 Step Response Estimates Using External Excitation, true step response (- -), estimates (-), true response (- -), 99% confidence bounds (-) .......... 30
3.7 Average Step Response Estimate Using External Excitation, estimate (-), true response (- -), 99% confidence bounds (-) ............... 31
3.8 Step Response Estimates Using External Excitation and Mismatched Prefilter, estimates (-), true response (- -), 99% confidence bounds (-) .......... 32
3.9 Average Step Response Estimate Using External Excitation and Mismatched Prefilter, estimate (-), true response (- -), 99% confidence bounds (-) .......... 33
3.10 Step Response Estimates Using External Excitation and Correct Prefilter, estimates (-), true response (- -), 99% confidence bounds (-) .......... 34
3.11 Average Step Response Estimate Using External Excitation and Correct Prefilter, estimate (-), true response (- -), 99% confidence bounds (-) .......... 34
3.12 Sample Excitation and Output Signals: Stationary Noise Case .......... 39
3.13 Average Noise Spectrum: Stationary Noise Case, true (o), actual (*) .......... 39
3.14 Sample Noise Signals: Stationary Noise Case, u(t) (- -), u(t) - G(z)K(z)S(z)u(t) (-) .......... 40
3.15 Sample Excitation and Output Signals: Non-Stationary Noise Case .......... 41
3.16 Average Noise Spectrum: Non-Stationary Noise Case, true (o), actual (*) .......... 41
4.1 Timeline of the Two Phase Experiment ......................... 47
4.2 Phase One Reference Signal .................................. 61
4.3 Phase One Input-Output Measurements .......................... 62
4.4 Phase Two Reference Signal .................................. 65
<table>
<thead>
<tr>
<th>Section Number</th>
<th>Section Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>Input-Output Measurements: Two Phase Approach</td>
<td>66</td>
</tr>
<tr>
<td>4.6</td>
<td>Input Signal FSF Energy Distribution, desired (.), actual (.)</td>
<td>67</td>
</tr>
<tr>
<td>4.7</td>
<td>FSF-Generated Step Response Estimate: Two Phase Approach, estimate (-), true response (-), 95% confidence bounds (- -)</td>
<td>67</td>
</tr>
<tr>
<td>4.8</td>
<td>Input-Output Measurements</td>
<td>69</td>
</tr>
<tr>
<td>4.9</td>
<td>FSF-Generated Step Response Estimate, estimate (-), true response (-), 95% confidence bounds (- -)</td>
<td>69</td>
</tr>
<tr>
<td>4.10</td>
<td>Input Signal FSF Energy Distribution</td>
<td>74</td>
</tr>
<tr>
<td>4.11</td>
<td>Average Input Signal FSF Energy Distribution, desired (.), average (.)</td>
<td>74</td>
</tr>
<tr>
<td>4.12</td>
<td>Fourier Amplitude Spectrum of the First Phase Reference Signal</td>
<td>76</td>
</tr>
<tr>
<td>4.13</td>
<td>Fourier Amplitude Spectrum of the Second Phase Reference Signal</td>
<td>77</td>
</tr>
<tr>
<td>4.14</td>
<td>Fourier Amplitude Spectrum of the Entire Reference Signal</td>
<td>77</td>
</tr>
<tr>
<td>4.15</td>
<td>Summation of Two Uncorrelated Signals</td>
<td>79</td>
</tr>
<tr>
<td>4.16</td>
<td>Magnitude of FSF Weighting Function, $W^0$</td>
<td>84</td>
</tr>
<tr>
<td>4.17</td>
<td>Magnitude of FSF Weighting Functions, $W^0$ (-), $W^1$ (-), $W^2$ (-)</td>
<td>85</td>
</tr>
</tbody>
</table>
Nomenclature

\( a(t) \) white noise sequence
\( A_k \) sine wave amplitude at the \( k \)th FSF frequency
\( b_i \) \( i \)th impulse response coefficient
\( d(t) \) dither signal at time, \( t \)
\( E \) expectation operator
\( F_k \) \( k \)th FSF filter
\( g(m) \) \( m \)th step response coefficient
\( G(z) \) discrete time process transfer function
\( G(e^{j \omega k}) \) discrete process frequency response at frequency \( \omega_k \)
\( H(z) \) noise model
\( J_M(\theta) \) least squares objective function
\( k \) FSF harmonic number
\( K(z) \) discrete process controller
\( M \) data length
\( M_P \) number of periods
\( n \) FSF model order
\( N \) process settling time, in units of sampling intervals
\( P \) period of excitation
\( Q(z) \) closed-loop transfer function from setpoint to input
\( r \) DFT harmonic number
\( S(z) \) sensitivity function
\( t \) time, in units of samples
\( T_s \) process settling time, in units of time
\( u(t) \) process input at time, \( t \)
\( v(t) \) disturbance at time, \( t \)
\( V \) disturbance vector
\( W^k \) \( k \)th FSF energy weighting function
\( W(k, m) \) step response weighting function
\( y(t) \) process output at time, \( t \)
\( y_r(t) \) reference signal at time, \( t \)
\( Y \) process output vector
\( z^{-1} \) backwards shift operator
Greek letters

\( \beta \)  \hspace{1cm} \text{autoregressive noise term} \\
\( \Delta t \)  \hspace{1cm} \text{sampling interval} \\
\( \epsilon(t) \)  \hspace{1cm} \text{residual at time, } t \\
\( \lambda_k \)  \hspace{1cm} k\text{th FSF spectrum energy} \\
\( \lambda_k' \)  \hspace{1cm} k\text{th optimal FSF spectrum energy after prefiltering} \\
\( \lambda_k'' \)  \hspace{1cm} k\text{th optimal FSF spectrum energy} \\
\( \lambda_k^{ref} \)  \hspace{1cm} k\text{th FSF spectrum energy of the reference signal} \\
\( \omega \)  \hspace{1cm} \text{frequency, in rad/sec} \\
\( \omega_k \)  \hspace{1cm} k\text{th FSF frequency, in rad} \\
\( \phi \)  \hspace{1cm} \text{FSF regressor vector (filtered input signal)} \\
\( \Phi(\omega) \)  \hspace{1cm} \text{power spectral density function} \\
\( \phi_k(t) \)  \hspace{1cm} k\text{th FSF filter output at time, } t \\
\( \sigma_a \)  \hspace{1cm} \text{standard deviation of the noise } a(t) \\
\( \tau_0 \)  \hspace{1cm} \text{fractional error in the process gain estimate} \\
\( \Theta \)  \hspace{1cm} \text{FSF parameter vector (frequency response parameters)}

Abbreviations

BLUE  \hspace{1cm} \text{Best Linear Unbiased Estimate} \\
DFT  \hspace{1cm} \text{Discrete Fourier Transform} \\
FFT  \hspace{1cm} \text{Fast Fourier Transform} \\
FIR  \hspace{1cm} \text{Finite Impulse Response} \\
FSF Model  \hspace{1cm} \text{Frequency Sampling Filter Model} \\
GLS  \hspace{1cm} \text{Generalized Least Squares} \\
IMC  \hspace{1cm} \text{Internal Model Control} \\
MIMO  \hspace{1cm} \text{Multi-Input Multi-Output} \\
PID Controller  \hspace{1cm} \text{Proportional Integral Derivative Controller} \\
PRBS  \hspace{1cm} \text{Pseudo Random Binary Signal} \\
PRESS  \hspace{1cm} \text{Prediction Error Sum of Squares} \\
SE  \hspace{1cm} \text{Sum of Errors} \\
SISO  \hspace{1cm} \text{Single-Input Single-Output} \\
SLS  \hspace{1cm} \text{Standard Least Squares}

Superscripts

\( \bar{x} \)  \hspace{1cm} \text{Desired value of } x \\
\( \hat{x} \)  \hspace{1cm} \text{Estimated value of } x \\
*  \hspace{1cm} \text{Complex conjugate transpose}
Chapter 1

Introduction

1.1 Motivation

The purpose of an identification experiment is to determine the dynamic characteristics of a given process. Input-output data collected from the experiment is used in an empirical model building exercise, where few or no assumptions are made regarding the structure of the model form. This is known as black box modelling. Ideally, the first step in the model building process should be experiment design. The aim here is selecting experimental conditions such that the data collected contains as much information as possible about the dynamical properties of the system that are to be identified. A carefully designed experiment will also be less costly, as shorter experiments and/or reduced output variations can be realized.

In many cases, the process to be identified is part of a closed-loop system. When safety is a major concern, or if the open-loop system is unstable, the regulator cannot be disabled during an identification experiment. These issues naturally lead to research into the identifiability of closed-loop systems. There has been an increasing amount of literature on the problem over the last twenty years, with perhaps the earliest being the work of Akaike (1967). Since then, much of the work has been aimed at developing identification algorithms that attempt to eliminate the effect of the feedback correlation in closed-loop data. These efforts have lead to the so-called indirect, joint input-output, and two-step methods. By breaking this inherent feedback correlation, these advanced algorithms allow the user to select an arbitrary bias distribution for the identified process model. That is, a prefilter designed to enhance certain frequency regions may be applied to the input-output
data. Such bias shaping is not possible with the direct approach, since the input-output data must be prefiltered with the inverse of the true noise model for consistent estimation of the process model.

Perhaps the most important parameter in an identification experiment is the power spectrum of the input signal. In a closed-loop system, the input is partly determined as feedback from other signals, thus complicating experiment design. This problem has been addressed by Forsell and Ljung (2000) and Zhu and van den Bosch (2000), but their results are theoretically complex, and largely depend on unknown entities (such as the system to be identified). While these authors present their designs in terms of optimal dither power spectrums (where the dither may be added to either the process input or setpoint), many practical considerations are not addressed. For example, what types of waveforms may be used to realize a desired power spectrum? How does one compute an excitation signal with a given power spectrum? In addition, the presence of a feedback mechanism presents a challenge for low frequency excitation. Integral action is used specifically to dampen out low frequency variations entering the process. In fact, with an integrator in the feedback loop, the sensitivity function has a frequency response of zero at the zero frequency. The use of a dither signal added to the controller output has been advocated by Forsell and Ljung (2000). However, with the transfer function from dither to the process input being the sensitivity function, it is impossible to excite the zero frequency in this manner. Some of these more practical issues will be addressed in this work.

In this thesis, the Frequency Sampling Filter (FSF) model is used to represent the process dynamics. While the FSF model has been used in the area of digital filter design for quite some time (Rabiner and Gold, 1975), the application of it in process identification is more recent (Goberdhansingh et al., 1992). Optimal experiment design for determining accurate step response models was shown to be relatively straightforward by Smektala (1998), where it was noted that inputs with a specific energy distribution will minimize the variance error in the FSF-generated step response estimate. This, along with some other unique properties of the FSF model, will be used to show that this model structure is particularly well-suited to studying the problem of optimal closed-loop identification.
1.2 Preliminaries

Figure 1.1 presents a schematic representation of a closed-loop system that will be considered throughout the thesis. The true open-loop system may be described by

\[ y(t) = G(z)u(t) + H(z)a(t) \]  

(1.1)

where \( u(t) \) is the discrete process input, \( y(t) \) is the discrete process output, and \( t \) is an integer variable used to enumerate the sampling instants. \( G(z) \) and \( H(z) \) are the unknown process and noise models, respectively. The variable \( z \) is the shift operator, such that \( z^{-1}u(t) = u(t - i) \), and \( a(t) \) is a zero mean white noise sequence with variance \( \sigma_a^2 \). The signal \( d(t) \) is a dither signal added to the controller output. With \( d(t) = 0 \), we may also write

\[ y(t) = G(z)K(z)S(z)y_{sp}(t) + H(z)S(z)a(t) \]  

(1.2)

\[ u(t) = S(z)K(z)y_{sp}(t) - H(z)S(z)K(z)a(t) \]  

(1.3)

where \( y_{sp}(t) \) is a reference signal, or setpoint which is uncorrelated with the noise \( a(t) \), and \( S(z) \) is known as the sensitivity function, given by

\[ S(z) = \frac{1}{1 + K(z)G(z)} \]  

(1.4)

The power spectrum of the input is

\[ \Phi_u(\omega) = |S(e^{j\omega})|^2 |K(e^{j\omega})|^2 \Phi_{y_{sp}}(\omega) + |S(e^{j\omega})|^2 |K(e^{j\omega})|^2 |H(e^{j\omega})|^2 \sigma_a^2 \]  

(1.5)

where \( \Phi_{y_{sp}}(\omega) \) is the spectrum of the reference signal. This expression may be broken into two terms, defined as follows:

\[ \Phi_u(\omega) = |S(e^{j\omega})|^2 |K(e^{j\omega})|^2 \Phi_{y_{sp}}(\omega) \]  

(1.6)

\[ \Phi_\epsilon(\omega) = |S(e^{j\omega})|^2 |K(e^{j\omega})|^2 |H(e^{j\omega})|^2 \sigma_a^2 \]  

(1.7)

It will be assumed that the single-input, single-output (SISO) process \( G(z) \) is stable, linear, and time-invariant. The reference signal \( y_{sp}(t) \) is assumed to be independent of the noise \( a(t) \). To avoid algebraic loops in the closed-loop system, it will also be assumed that either \( G(z) \) or \( K(z) \) contains at least one unit of delay.
1.3 Literature Review

Several different approaches to closed-loop identification of the unknown process model $G(z)$ have been presented in the literature. This section begins with a review of these approaches, restricting the discussion to those that may be classified as prediction error methods. Also included in this section is a review of the existing literature on experiment design for closed-loop identification experiments.

1.3.1 Closed-Loop System Identification

There are three main variations of the prediction error method that have been developed for closed-loop identification; these are known as the direct approach, the indirect approach, and the joint input-output approach. Each of these has its own advantages, and they may be distinguished by the assumptions on which they are based.

The direct approach is the straightforward application of a prediction error method to the input-output data. No assumptions about the nature of the feedback are made, and dither or setpoint signals are ignored, even if measured. Forssell and Ljung (1999) note that the direct approach gives optimal accuracy, but requires a correct noise model to avoid bias in the process model estimate (see Chapter 3 for a more detailed analysis). The most significant consequence of this is that the user cannot specify an arbitrary frequency domain weighting function to shape the bias distribution of the estimated process model.
The general assumption made when applying the \textit{indirect approach} is that the feedback mechanism is known. This does not only include the controller, which may be nonlinear, but also the anti-windup functions and other possible nonlinearities in the loop. This is therefore a rather restrictive assumption. The indirect approach also requires the measurement of an external signal such as a dither or setpoint. Given this information, the idea behind the indirect approach is to first identify the closed-loop system, then compute the open-loop process model. For example, referring to Figure 1.1, assume that $y_{sp}(t) = 0$. The closed-loop system then becomes

$$ y(t) = G_c(z)d(t) + H_c(z)a(t) $$

where $G_c(z) = G(z)S(z)$ and $H_c(z) = H(z)S(z)$. Since $d(t)$ and $a(t)$ are uncorrelated, the function $G_c(z)$ can be estimated in an open-loop fashion. This means that if the process model and noise model are independently parameterized, it is possible to consistently estimate $\hat{G}_c(z)$, irrespective of the noise contribution (Ljung, 1999). An estimate of the open-loop process model $\hat{G}(z)$ can then be computed by solving the equation

$$ \hat{G}(z) = \frac{\hat{G}_c(z)}{1 - \hat{G}_c(z)K(z)} $$

The difficulty with this type of solution is that the calculated $\hat{G}(z)$ will be of high order. This can be avoided if Equation 1.9 is parameterized as (van Donkelaar and van den Hof, 1997)

$$ \hat{G}(z, \theta) = \frac{\hat{G}_c(z, \theta)}{1 - \hat{G}_c(z, \theta)K(z)} $$

In this case, the open-loop and closed-loop systems are parameterized within the same model structure (e.g. ARMAX or Box-Jenkins). Algebraic and numeric problems are thus avoided. To outline the idea, consider indirect identification using ARMAX models. Factorize $K(z)$ as

$$ K(z) = \frac{X(z)}{Y(z)} $$

where the polynomials $X(z)$ and $Y(z)$ are coprime. Let the closed-loop and open-loop model structures be

$$ A_c(z)y(t) = B_c(z)d(t) + C_c(z)a(t) $$

$$ A(z)y(t) = B(z)u(t) + C(z)a(t) $$
respectively. Since

\[ G_c = \frac{G(z)}{1 + G(z)K(z)} \]  \hspace{1cm} (1.14)

and

\[ H_c = \frac{H(z)}{1 + G(z)K(z)} \]  \hspace{1cm} (1.15)

it follows that the open-loop polynomials may be solved from

\begin{align*}
A_c(z) &= A(z)Y(z) + B(z)X(z) \\
B_c(z) &= B(z)Y(z) \\
C_c(z) &= C(z)Y(z)
\end{align*}  \hspace{1cm} (1.16)

With the polynomials \( A_c(z), B_c(z), \) and \( C_c(z) \) being estimated from a closed-loop experiment, Equation 1.16 represents a system of linear equations in the open-loop parameters.

**Joint input-output** methods exploit the fact that the output and input variables may be expressed in terms of the external signals \( d(t) \) and \( a(t) \), as in Equation 1.8 for the output, and

\[ u(t) = S(z)d(t) - H(z)S(z)K(z)a(t) \]  \hspace{1cm} (1.17)

for the input. Thus, the closed-loop identification problem can be broken into two open-loop problems (this is commonly referred to as the two-step approach).

In the first step, measurements of \( u(t) \) and \( d(t) \) are used to identify the sensitivity function, \( S(z) \) (Equation 1.17). As in the indirect approach, this first step is an open-loop problem.

For the second step, van den Hof and Schrama (1993) suggest an approach where the input signal is reconstructed, using the estimate \( \hat{S}(z) \) as follows:

\[ \hat{u}(t) = \hat{S}(z)d(t) \]  \hspace{1cm} (1.18)

Substituting this into Equation 1.8 gives

\[ y(t) = G(z)\hat{u}(t) + H(z)S(z)a(t) \]  \hspace{1cm} (1.19)

It follows that \( G(z) \) can also be estimated in an open-loop fashion, since \( \hat{u}(t) \) and \( a(t) \) are uncorrelated. The quality of the final process model \( \hat{G}(z) \) will naturally depend on the accuracy of the identified sensitivity function. The ability to use an arbitrary prefilter on
the input-output data does however allow the user to specify frequency regions where the modelling error should be small.

In an alternative formulation, Huang and Shah (1997) use the estimated sensitivity function \( \tilde{S}(z) \) as a prefilter on the output data. Equation 1.8 becomes

\[
\tilde{S}^{-1}(z)y(t) = y_f(t) = \frac{S(z)}{\tilde{S}(z)}G(z)d(t) + \frac{S(z)}{\tilde{S}(z)}H(z)a(t)
\] (1.20)

The variables \( y_f(t) \) and \( d(t) \) can then be used to obtain an estimate of the process, \( G(z) \), in an open-loop fashion. This method is similar to the one proposed by van den Hoel and Schrama (1993) in that the accuracy of the estimated sensitivity function is a main issue.

The asymptotic variance of estimates obtained from the direct approach meets the Cramèr-Rao lower limit (Ljung, 1999), meaning that the direct method is optimal with respect to accuracy. The indirect methods generally give suboptimal accuracy, and cannot be applied to systems with unknown feedback mechanisms. The statistical properties and overall applicability of the direct method make it the preferred approach to closed-loop identification, an opinion expressed by Forssell and Ljung (1999) and Esmaili et al. (2000). Based on these references, direct identification is used in this thesis.

1.3.2 Experimental Design

Experiment design involves the selection of conditions such that the experiment can be maximally informative at a given cost. In system identification, these 'conditions' include things such as, but not limited to, the frequency content of excitation signals, the duration of the experiment, the type of waveform used, and data prefilters. The 'costs' are largely financial in nature, since identification experiments require perturbation of a plant, thus affecting its output performance. The design of an experiment must also take into consideration the intended application of the results, the extent of prior knowledge, and constraints on the system.

The problem of optimal experiment design for closed-loop identification has been addressed recently by Forssell and Ljung (2000). The authors give expressions for the optimal controller and optimal test signal that will minimize the variance error in the identified process model, under constraints on a linear combination of the input and output variances. For the case where only the misfit in \( \tilde{G}(z) \) is penalized, the optimal controller and test
signal spectrum under the constraint
\[ \int_{-\pi}^{\pi} a\Phi_u + b\Phi_y \, d\omega \leq 1 \quad (1.21) \]
are given by
\[ K_{opt} = \arg \min_K \int_{-\pi}^{\pi} \frac{a|K|^2 + b\Phi_v}{|1 + GK|^2} \, d\omega \quad (1.22) \]
\[ \Phi_d^{opt} = \mu \sqrt{\Phi_v Q} \frac{1 + GK_{opt}}{\sqrt{\alpha + \beta |G|^2}} \quad (1.23) \]
where \( \mu \) is a constant, adjusted so that
\[ \int_{-\pi}^{\pi} a\Phi_u + b\Phi_y \, d\omega = 1 \quad (1.24) \]
and \( \Phi_v = |H|^2 \sigma_v^2 \) is the noise spectrum. The weighting function \( Q \) is selected by the user, and reflects the relative importance of a good fit in \( \hat{G} \) for different frequencies.

The authors acknowledge that the main drawback with this solution is that it depends on the unknown system \( G(z) \). However, they propose the use of iterations as a way around this, where the ‘true’ unknown system is replaced by a current estimate of the system. It should also be noted that this solution suggests that as long as there is a constraint on the output variance, closed-loop experiments are optimal. This observation is supported by Esmaili et al. (2000). Gevers et al. (2001) and Hjalmarsson et al. (1996) also show that closed-loop experiments are preferred when the objective of the identification is model-based control design.

Zhu and van den Bosch (2000) argue that the above results are too complex and lead to infeasible designs. Founded on the same principles, the authors offer an alternative optimal design for systems using Internal Model Control (IMC) for both identification and control. The design has also been simplified by assuming that only one signal is constrained, in contrast to Equation 1.21. The optimal test signal in this case is related to the error of the current model \( (G - \hat{G}) \). Therefore, assuming the error in the current model estimate is small, the dependence of the optimal design on the unknown process should be negligible.

The ideas presented in the above two papers differ significantly. While the work of Forssell and Ljung (2000) has a strong theoretical foundation and is somewhat complicated, the design by Zhu and van den Bosch (2000) though simpler, applies only to systems using IMC. This thesis will attempt to bridge the gap between these two extremes, providing insight into practical issues regarding optimal identification experiments.
1.4 Outline of the Thesis

Chapter 2 introduces the Frequency Sampling Filter (FSF) model structure. The FSF model is a linear transformation of the $N$th order, finite impulse response (FIR) model. A particularly interesting property of the FSF model structure is that the model order can be reduced to $n$, which is generally much less than $N$. With the reduction of model order, the correlation matrix associated with the least squares estimate of the FSF model is generally better conditioned than that of the equivalent FIR model. FSF-generated step response estimates are therefore smoother than those estimated from an FIR model. An important property of the FSF model structure for the work in this thesis is that the correlation matrix offers an indication of the quality of input excitation. It is shown that a particular distribution of input energy minimizes the variance in the FSF-generated step response estimate, forming the basis for the optimal design presented in this thesis. Chapter 2 also describes how an input signal consisting of a sum of sinusoids can be designed to achieve a desired energy distribution.

In Chapter 3, the role of the noise model and data prefilters in closed-loop identification is examined. The noise model has a direct effect on the bias and variance of process models estimated from closed-loop data. Expressions are given to support the well-known fact that when using the direct approach, a consistent estimate of the process model can be obtained only when the input-output data is prefiltered with the inverse of the true noise model. Chapter 3 also provides some insight into the value of that part of the input signal which is generated from the feedback. A series of MonteCarlo simulations are provided to illustrate the effects of external excitation and data prefiltering on the bias and variance error of FSF-generated step response estimates. The remainder of the chapter highlights the use of periodic excitation in identification experiments. A simple averaging technique may be used to reduce the data set into a single period. It is shown that the noise properties can then be recovered from the data set, allowing for estimation of a noise model.

Chapter 4 represents the major contribution of this thesis to the area of experiment design. A new two-phase approach to the problem of optimal experiment design for closed-loop experiments is proposed. The objective of this design is to minimize the variance error in step response models estimated from closed-loop data. An expression for the optimal
energy distribution in the reference signal is developed. The expression is similar to the optimal designs presented in the literature, in that it is a function of unknown parameters. In the proposed two-phase approach, the aforementioned unknown parameters are estimated during the first phase of the experiment, allowing the optimal design to be realized over the combined two phases. Significant consideration is given to practical aspects relating to the design of each phase of the experiment. In particular, the use of periodic excitation is advocated as a means of improving the computational efficiency of the proposed two-phase approach. Simulation studies have been included to illustrate the practical use of the two-phase approach.
Chapter 2

Frequency Sampling Filter Model Identification

The FSF model is well known in the area of digital filter design (Rabiner and Gold, 1975), and was first introduced to system identification and automatic control by Bitmead and Anderson (1981). The FSF model is a linear transformation of the $N$th order, finite impulse response (FIR) model and, as such, is represented by $N$ frequency response parameters. The FIR model is widely used in the field of process identification because it requires no prior knowledge of the process other than its settling time. The FSF model and some details about the FSF identification algorithm relevant to this thesis are given in this chapter.

2.1 Frequency Sampling Filter (FSF) Model Derivation

Consider a single-input, single-output (SISO) process that is stable, linear and time-invariant and can be represented by the discrete-time FIR model

$$G(z) = \sum_{i=0}^{N-1} b_i z^{-i}$$  \hspace{1cm} (2.1)

The model given in Equation 2.1 is characterized by the impulse response coefficients $b_i$, $i = 0, ..., N - 1$ and model order $N$. Note that $b_i \approx 0$ for all $i \geq N$. The model order, $N$, is usually chosen such that $N \approx \frac{T_s}{\Delta t}$, where $\Delta t$ is the sampling interval and $T_s$ is the settling time.

The FSF model is derived by substituting the following inverse discrete Fourier transform (DFT) into Equation 2.1

$$b_i = \frac{1}{N} \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} G(e^{j \frac{2\pi k}{N}}) e^{j \frac{2\pi i k}{N}}$$  \hspace{1cm} (2.2)
Interchanging the summations and noting that
\[
\sum_{i=0}^{N-1} e^{j2\pi ik/N} z^{-i} = \frac{1 - z^{-N}}{1 - e^{j2\pi k/N} z^{-1}}
\]
we can describe the FSF model as
\[
G(z) = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} G(e^{j\omega_k}) \frac{1 - z^{-N}}{N 1 - e^{j\omega_k} z^{-1}}
\]
where \( \omega_k = \frac{2\pi k}{N} \) radians, \( k = 0, \pm 1, \ldots, \pm \frac{N-1}{2} \) and these are called the FSF frequencies. A general process model that describes the relationship between the input signal, \( u(t) \), and the process output, \( y(t) \), is therefore
\[
y(t) = G(z)u(t) = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} G(e^{j\omega_k}) \frac{1 - z^{-N}}{N 1 - e^{j\omega_k} z^{-1}} u(t)
\]

The schematic diagram of the FSF model structure, in Figure 2.1, illustrates how the input signal, \( u(t) \), passes through \( N \) parallel band pass filters and the filter outputs \( \phi^k(t) \) are weighted by the process frequency response, \( G(e^{j\omega_k}) \), corresponding to each filter's centre frequency, \( \omega_k \). The sum of these weighted frequency components then forms the process output, \( y(t) \).
The FSF model has the following properties:

- With fast sampling, the FSF model parameters converge to their continuous-time counterparts at $\omega = 0, \frac{2\pi}{T_s}, \ldots, \frac{\pi}{T_s}$ for a fixed $T_s$. As the sampling interval $\Delta t$ decreases, the number of parameters associated with the model increases but only in the high frequency region.

- Based on the above property, there exists an odd integer $n$ such that for all $k$ where $\frac{n-1}{2} < |k| \leq \frac{N-1}{2}$, the magnitudes of the FSF model parameters, $G(e^{j\omega})$, are approximately equal to zero; this is limited to processes that can be described by strictly proper transfer functions. As $\Delta t \to 0$, $n$ becomes independent of the choice of sampling interval and is called the reduced model order which is generally much less than $N$. The reduced order FSF model is written as follows:

$$G(z) \approx \sum_{k=-\frac{n-1}{2}}^{\frac{n-1}{2}} G(e^{j\omega_k}) \frac{1}{N} \frac{1 - z^{-N}}{1 - e^{j\omega_k} z^{-1}}$$

with the assumption that $G(e^{j\omega_k})$ is negligible for $\frac{n-1}{2} < |k| \leq \frac{N-1}{2}$. The issue of how to determine the reduced model order, $n$, is addressed by Wang and Cluett (1996).

- Due to the reduction of model order, the correlation matrix associated with the least squares estimate of the FSF model is generally better conditioned than that of the equivalent FIR model. The correlation matrix associated with the FIR model is only well-conditioned when the periodogram of the input signal is approximately equal at all frequencies. However, since input signals with only low and medium frequency content are typically used in the process industries, the correlation matrix for the FIR model will almost always be ill-conditioned. A singular or nearly-singular correlation matrix will lead to inaccurate model parameter estimates. The FSF model has the advantage in this respect, mainly because of the reduction in model order, by including only those parameters corresponding to $0 < |k| \leq \frac{n-1}{2}$, i.e. frequencies in the low to medium range.

### 2.2 Formulation of the Least Squares Problem

Consider again Figure 2.1, the block diagram of the FSF model structure. This schematic shows that the outputs of the FSF filters at any sampling instant, $t$, form the elements in
the FSF regressor vector, \( \phi(t) \), defined as

\[
\phi(t) = \begin{bmatrix}
\phi^0(t) & \phi^1(t) & \phi^{-1}(t) & \ldots & \phi^{\frac{n-1}{2}}(t) & \phi^{-\frac{n-1}{2}}(t)
\end{bmatrix}^T
\]  
(2.7)

These elements represent the FSF frequency components present in the input signal. More specifically,

\[
\phi^k(t) = F^k(z)u(t)
\]  
(2.8)

where \( \phi^k(t) \) represents the output of the \( k \)th FSF filter, \( F^k(z) \):

\[
F^k(z) = \frac{1}{N} \frac{1 - z^{-N}}{1 - e^{j\omega_k} z^{-1}}
\]  
(2.9)

\[
F^k(z) = \frac{1}{N} \left(1 + e^{j\omega_k} z^{-1} + e^{2j\omega_k} z^{-2} + \ldots + e^{(N-1)j\omega_k} z^{-(N-1)}\right)
\]  
(2.10)

A more general process model is now given by

\[
y(t) = G(z)u(t) + v(t)
\]  
(2.11)

where \( v(t) \) is the output disturbance sequence given by \( v(t) = H(z)a(t) \), and where \( H(z) \) is the noise model and \( a(t) \) is a white noise sequence. Combining Equations 2.6, 2.8, and 2.11 yields

\[
y(t) = \sum_{k=-\frac{n-1}{2}}^{\frac{n-1}{2}} G(e^{j\omega_k})\phi^k(t) + v(t)
\]  
(2.12)

Equation 2.12 can also be expressed in matrix form, for \( M \) measurements, as:

\[
Y = \Phi \Theta + V
\]  
(2.13)

where,

\[
Y = \begin{bmatrix}
y(0) & y(1) & \ldots & y(M-1)
\end{bmatrix}^T
\]  
(2.14)

\[
V = \begin{bmatrix}
v(0) & v(1) & \ldots & v(M-1)
\end{bmatrix}^T
\]  
(2.15)

\[
\Phi = \begin{bmatrix}
\phi^0(0) & \phi^1(0) & \phi^{-1}(0) & \ldots & \phi^{\frac{n-1}{2}}(0) & \phi^{-\frac{n-1}{2}}(0) \\
\phi^0(1) & \phi^1(1) & \phi^{-1}(1) & \ldots & \phi^{\frac{n-1}{2}}(1) & \phi^{-\frac{n-1}{2}}(1) \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\phi^0(M-1) & \phi^1(M-1) & \phi^{-1}(M-1) & \ldots & \phi^{\frac{n-1}{2}}(M-1) & \phi^{-\frac{n-1}{2}}(M-1)
\end{bmatrix}
\]  
(2.16)

and

\[
\Theta = \begin{bmatrix}
G(e^{j\omega_0}) & G(e^{j\omega_1}) & G(e^{-j\omega_1}) & \ldots & G(e^{j\omega_{\frac{n-1}{2}}}) & G(e^{-j\omega_{\frac{n-1}{2}}})
\end{bmatrix}^T
\]  
(2.17)
Equation 2.13 is solved for $\Theta$ by minimizing the sum of squared prediction errors (i.e. the least squares solution):

$$\hat{\Theta} = (\Phi^*\Phi)^{-1}\Phi^*Y$$  

(2.18)

An estimate of the noise variance $\sigma_n^2$ is given by

$$\sigma_n^2 = \frac{(Y - \Phi\Theta)^*(Y - \Phi\Theta)}{M + 1 - n}$$  

(2.19)

If the noise is not white, then the estimate $\hat{\Theta}$ from Equation 2.18 will not be the best linear unbiased estimate (BLUE). However, with prior knowledge of the noise model, the input-output data may be pre-whitened in order to obtain the BLUE of $\Theta$.

### 2.3 Step Response Models

A step response model can be generated from the FSF model by relating the coefficients of the FSF model to the coefficients of the step response model. Let $g(m)$ denote the step response coefficient at sampling instant $m$, where $m = 0, 1, \ldots, N - 1$. The step response coefficients, $g(m)$, are related to the impulse response coefficients, $b_i$, through

$$g(m) = \sum_{i=0}^{m} b_i$$  

(2.20)

Substituting the reduced order version of Equation 2.2 into Equation 2.20 yields

$$g(m) = \frac{1}{N} \sum_{i=0}^{m} \sum_{k=-N/2}^{N/2} G(e^{j\omega k})e^{j\omega ki}$$  

(2.21)

Interchanging the summations and noting Equation 2.3, gives

$$g(m) = \sum_{k=-N/2}^{N/2} G(e^{j\omega k})w(k, m)$$  

(2.22)

where

$$w(k, m) = \frac{1}{N} \frac{1 - e^{j\omega k}(m + 1)}{1 - e^{j\omega k}}$$  

(2.23)

and $w(k, m)$ is called the step response weighting function. Equation 2.22 gives the explicit relationship between the coefficients of the FSF model and the coefficients of the step response model.

Wang and Cluett (1997) have shown that reduction of the FSF model order from $N$ to $n$ is justified by examining the contribution of the high frequency parameters to the overall step
response model. For $k = 0$, the step response weighting function, $w(0, m)$, linearly increases to unity as $m = 0, 1, \ldots, (N - 1)$. For $k > 0$, the step response weighting function, $w(k, m)$, can be broken down into two separate terms that operate on the real and imaginary parts of $G(e^{jk\omega})$. In general, the magnitude of both the real and imaginary weighting functions decreases as the frequency corresponding to $k$ increases. Therefore, $G(e^{jk\omega})$ is the most important parameter, and subsequent frequency response parameters become less and less significant in terms of their contribution to the step response model. The fact that the step response weighting functions de-emphasize the high frequency parameters allows for a significant reduction in the FSF model order without much loss of accuracy in the step response model.

2.4 Confidence Bounds

There are two major sources that contribute to uncertainty in an estimated model. The first is the presence of unmodelled dynamics, the bias error, and the second is the presence of disturbances in the process output, the variance error. It is reasonable to assume that the effect of unmodelled dynamics is secondary to the effect of disturbances, especially when dealing with finite data lengths and limited input amplitudes. The FSF model confidence bounds are therefore a quantification of the effect of disturbances on model uncertainty. The FSF model confidence bounds are obtained via the use of statistical methods which result in 'soft bounds' or probabilistic bounds on model uncertainty. The confidence bounds for the FSF model and the FSF-generated step response model are discussed in what follows.

The following assumptions apply to the FSF model confidence bounds and the FSF-generated step response model confidence bounds:

1. The process has finite settling time $T_s$ and the parameter $N$ is chosen to be greater than or equal to $\frac{T_s}{\Delta}$ such that the neglected dynamics do not produce any significant bias on the parameter estimates.

2. The disturbance $v(t)$ is a zero mean, white and normally distributed random sequence with variance $\sigma_v^2$.

3. $n = N$, or $n$ is chosen such that the neglected frequency parameters are sufficiently small.
2.4.1 FSF Model Confidence Bounds

Since the FSF model is obtained via the least-squares estimator, the covariance of the estimate is given by

\[ E[(\hat{\Theta} - \Theta)(\hat{\Theta} - \Theta)^\top] = (\Phi^\top \Phi)^{-1} \sigma_a^2 \]  

(2.24)

The probabilistic bound on the additive modelling error between the true process frequency response and the estimated process frequency response, obtained via the least squares solution, is given by

\[ |\Delta G(z)|_{z=e^{j\omega_k}} = |\hat{G}(z) - G(z)|_{z=e^{j\omega_k}} \leq p(c_i \sigma_a^2)^{1/2} \]  

(2.25)

where \( c_i \) denotes the \( i \)th diagonal element of \((\Phi^\top \Phi)^{-1}\). The parameter \( p \) is a positive, real number which determines the probability of the true modelling error satisfying the bound, e.g. \( p = 1, 2, 3, \ldots \) means that the modelling error satisfies the bound with probability \(0.683, 0.954, 0.997, \ldots\) according to the specified level of the normal distribution (Goberdhan Singh et al., 1992). The right-hand side of the inequality in Equation 2.25 is the radius of the circular confidence bound centered on the FSF model parameter \( G(e^{j\omega_k}) \).

2.4.2 FSF-generated Step Response Model Confidence Bounds

The step response confidence bound is derived by first representing the step response coefficients as a linear transformation of the estimated FSF model parameters and then mapping the covariance matrix from the frequency domain to the time domain.

Equation 2.22 can be rewritten as

\[ \hat{g}(m) = W(m) \hat{\Theta} \]  

(2.26)

where \( W(m) = \begin{bmatrix} w(0, m) & w(1, m) & w(-1, m) & \ldots & w(n-1, m) & w(n-1, m) \end{bmatrix} \) with \( w(k, m) \) defined in Equation 2.23. Now the error, at the \( m \)th sampling instant, between the true process step response weight, \( g(m) \), and the estimated step response weight, \( \hat{g}(m) \), is bounded by

\[ |\hat{g}(m) - g(m)| \leq p\delta(m) \]  

(2.27)

where \( \delta(m)^2 \), the variance of the step response estimate at the sampling instant \( m \), is given by

\[ \delta(m)^2 = W(m)(\Phi^\top \Phi)^{-1}W(m)^* \sigma_a^2 \]  

(2.28)
The trajectory of the true step response, \( g(m) \), for \( m = 0, 1, \ldots, N - 1 \) lies inside the envelope generated by \( \hat{g}(m) \pm p\delta(m) \) with the probability related to \( p \). This envelope provides a confidence bound for the FSF-generated step response model.

### 2.5 Optimal FSF Input Spectrum

A unique property of the FSF correlation matrix, \( \Phi^*\Phi \) which appears in Equation 2.18, is that it offers an indication of the quality of the input excitation in the vicinity of the FSF frequencies. The FSF correlation matrix is given by:

\[
\Phi^*\Phi = \begin{bmatrix}
\nu_{11} & \nu_{12} & \nu_{13} & \cdots & \nu_{1(n-1)} & \nu_{1n} \\
\nu_{21} & \nu_{22} & \nu_{23} & \cdots & \nu_{2(n-1)} & \nu_{2n} \\
\nu_{31} & \nu_{32} & \nu_{33} & \cdots & \nu_{3(n-1)} & \nu_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\nu_{n1} & \nu_{n2} & \nu_{n3} & \cdots & \nu_{n(n-1)} & \nu_{nn}
\end{bmatrix}
\]

where \( \nu_{ij} = \sum_{t=0}^{M-1} \phi^k(t)\phi^q(t) \). Note that \( i \) and \( j \) are integers in the range \([1, n]\), and \( k \) and \( q \) are integers in the range \([-\frac{n-1}{2}, \frac{n-1}{2}]\).

Wang and Cluett (1997) show that the elements of the correlation matrix are weighted sums of the energy contributions from different frequencies in the input signal. Of particular importance, the diagonal elements of the correlation matrix are representative of the amount of input signal energy at the FSF frequencies because they are proportional to the periodogram, \( |U(e^{j\omega_k})|^2 \), of the input signal in the vicinity of \( \omega_k \); e.g.

\[
\nu_{kk}^k = \sum_{t=0}^{M-1} \phi^k(t)\phi^k(t) \approx \frac{M}{N} |U(e^{j\omega_k})|^2
\]

where

\[
U(e^{j\omega_k}) = \frac{1}{M} \sum_{t=0}^{M-1} u(t)e^{-j2\pi kt/N}
\]

Since these diagonal elements provide an approximation of the distribution of input signal energy at the FSF frequencies, they will be called the ‘FSF spectrum energies’. The distribution of these FSF spectrum energies is called the ‘FSF spectrum shape’ and is given...
by

$$\zeta_k = \frac{\lambda_k}{\sum_{l=0}^{n-1} \lambda_l} \times 100 \text{ for positive } k$$  \hspace{1cm} (2.32)

where $\lambda_k = \nu^{k,k}$. The quantity, $\zeta_k$, is interpreted as the percentage of energy at the FSF frequency $\omega_k, k = 0, \ldots, \frac{n-1}{2}$. As well, the ‘FSF spectrum ratios’ are defined as:

$$\psi_r = \frac{\zeta_0}{\zeta_r} = \frac{\lambda_0}{\lambda_r}$$  \hspace{1cm} (2.33)

where $r = 1, 2, \ldots, \frac{n-1}{2}$. The FSF spectrum ratio $\psi_r$ is interpreted as the proportion of energy at the FSF frequency $\omega_r, r = 1, 2, \ldots, \frac{n-1}{2}$, relative to the energy at the zero frequency. It should be noted that the FSF spectrum energies $\nu^{k,k}$ and $\nu^{-k,-k}$ correspond to the positive and negative FSF frequency $\omega_k$ and are therefore equal.

The link between the diagonal elements of the correlation matrix and the energy content of the input signal naturally led to the work of Smektala (1998) in which an ‘optimal’ FSF spectrum shape was defined. Since the quality of a FSF-generated step response model can be assessed on the basis of the area enclosed by the model’s confidence bounds, the following objective function was defined:

$$SE = \sum_{m=0}^{N-1} 2\delta(m)$$  \hspace{1cm} (2.34)

where $SE$ stands for the sum of errors, $\delta(m)$ is the standard deviation of the FSF-generated step response model parameter $\tilde{g}(m)$ ($\delta(m)^2$ is defined in Equation 2.28), and the number ‘2’ indicates a 95% confidence bound.

The optimal FSF spectrum ratios were found in Smektala (1998) by minimization of this objective function using a direct search optimization routine: the result for the case when $n = 13$ is presented in Table 2.1. So, for a fixed level of total input energy, this distribution will minimize the uncertainty in the resulting step response estimate. While the optimization was based on the assumption of a diagonal correlation matrix, Smektala (1998) suggests that this assumption can be relaxed to requiring that the correlation matrix is diagonally-dominant. Unlike the optimal designs presented by authors such as Forssell

<table>
<thead>
<tr>
<th>$\psi_1^{\text{optimal}}$</th>
<th>$\psi_2^{\text{optimal}}$</th>
<th>$\psi_3^{\text{optimal}}$</th>
<th>$\psi_4^{\text{optimal}}$</th>
<th>$\psi_5^{\text{optimal}}$</th>
<th>$\psi_6^{\text{optimal}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.7</td>
<td>3.4</td>
<td>5.0</td>
<td>6.6</td>
<td>8.2</td>
<td>9.7</td>
</tr>
</tbody>
</table>

Table 2.1: Optimal FSF Spectrum Ratios
and Ljung (2000) for example, this result is process independent, with only an estimate of the settling time \( N \) being needed \textit{a priori}. It should be noted that the result presented in Table 2.1 applies only to systems corrupted by white noise. In cases where the noise in the system is not white, prior knowledge of the noise model may be used to pre-whiten the input-output data. The objective function given by Equation 2.34 will then be minimized if the energy in the pre-whitened input signal has the distribution defined in Table 2.1.

Smektala (1998) pointed out the optimization problem can be resolved for a larger correlation matrix, i.e. for more FSF frequencies, but it was found that frequencies greater than \( \omega_b \) do not contribute significantly to the magnitude of the \( SE \). This thesis will adopt the same objective function given in Equation 2.34 as a measure of the estimated model quality, and unless otherwise stated, the reduced FSF model order \( (n) \) will be fixed at 13.

### 2.6 Multisine Design for FSF Model Identification

Under open-loop conditions, an input signal consisting of a linear combination of sinusoids is a particularly efficient way of achieving a desired FSF spectrum. Goberthansingh et al. (1992) show that a multisine with frequencies corresponding to the FSF frequencies produces orthogonal FSF filter outputs. That is, for an input signal of the form

\[
u(t) = \sum_{k=0}^{n-1} A_k \cos \left( \frac{2\pi kt}{N} \right)
\]  

(2.35)

the diagonal elements of the correlation matrix are given by

\[
\lambda_k = \frac{A_k^2 M}{4}
\]  

(2.36)

where \( M \) is the length of the input signal, and is chosen to be an integer multiple of \( N \). So for a given set of desired FSF spectrum energies, \( \lambda_k \), the amplitudes \( A_k \) can be determined \textit{a priori} via Equation 2.36. For this type of input, any off-diagonal elements of the correlation matrix are exactly zero. The problem with using a multisine with the phases of each component chosen arbitrarily is that the signal may exhibit sharp peaks, i.e. it may have a high peak factor.

An algorithm proposed by van den Bos and Krol (1979), referred to here as the van den Bos algorithm, may be used to construct test signals for system identification. This is an iterative algorithm which produces a binary signal that best approximates, in a least squares
sense, a specified Fourier amplitude spectrum. An intermediate step in each iteration of this algorithm is the construction of a multisine that has an amplitude spectrum exactly equal to the desired one. Convergence of this algorithm occurs when the multisine and binary signals have equal phases at the frequencies of interest. Based on the observation that binary signals possess good peak factors, it is reasonable to assume that the multisine constructed by the van den Bos algorithm will have a low peak factor as well. This algorithm has been used for the multisine input designs in this thesis.

The use of the van den Bos algorithm to achieve a specified FSF spectrum consists of the following steps:

1. Define the desired FSF spectrum, \( \overline{\lambda}_k \), for \( k = 0, \ldots, \frac{n-1}{2} \)

2. Choose the period of the multisine, \( P \) (an integer number of sampling instants), and the total length of the input sequence, \( M \) (an integer multiple of the period length, \( P \)). The period and total length should be integer multiples of the process settling time \( N \).

3. Calculate the required Fourier amplitude spectrum, \( U(e^{j\omega_r}) \), where \( \omega_r = \frac{2\pi r}{P} \) for \( r = 0, \ldots, P - 1 \). The Fourier amplitude spectrum is a vector of length \( P \), with all elements being equal to zero, except for the relatively few that correspond to FSF frequencies. The non-zero elements are given by

\[
U\left(\frac{P k}{N}\right) = \sqrt{\frac{\overline{\lambda}_k}{M}}
\]  

(2.37)

for \( k = 0, \ldots, \frac{n-1}{2} \). See Appendix A for a derivation of this result.

4. Use the van den Bos technique to calculate one period of the multisine. The complete input sequence, composed of \( M/P \) periods, will have the specified FSF spectrum and a good peak factor.

It should be noted that any routine for multisine synthesis may be used as an alternative to the one chosen here. For example, see Schroeder (1970) and van der Ouderaa et al. (1988). Most are similar to that of van den Bos, in that they use a swapping technique between time and frequency domains. Therefore, Steps 1 to 3 will apply in any case, since these simply establish the link between the FSF spectrum and the Fourier amplitude spectrum.

\(^1\)Also denoted as \( U(r) \).
2.7 Summary

This chapter has reviewed the development of the FSF model. In Section 2.1, the FSF model was derived from a linear transformation of the popular FIR model. In Section 2.2, the FSF model is combined with a standard least squares estimator in order to obtain a frequency domain model of the process. Section 2.3 shows how a step response model can be generated from the FSF model. In addition, probabilistic bounds on the modelling error in the frequency and time domain models can be obtained, as outlined in Section 2.4.

In Section 2.5, a crucial link between the input signal and the FSF correlation matrix was established. The diagonal elements of the correlation matrix were shown to be a measure of the input signal energy at the FSF frequencies, and are therefore referred to as the ‘FSF spectrum energies’. With a certain distribution of the FSF spectrum energies, i.e. the optimal FSF spectrum shape, the uncertainty in the step response model is minimized.

Input signal design for FSF model identification is introduced in Section 2.6. It was suggested that an input signal consisting of a linear combination of sinusoids is the most efficient type of design for achieving a desired FSF spectrum. Multisines constructed by algorithms such as the one proposed by van den Bos have the added benefit of exhibiting favourable peak factors.
Chapter 3

The Role of the Noise Model

In this chapter, the role of the noise model and prefilters when using prediction error methods in closed-loop identification is examined. The similarity between the prefilter and the noise model, and how they relate to the asymptotic bias and variance of the identified process model, is considered in Section 3.1. Simulation studies are used to illustrate some of these concepts, and these are included in Section 3.2. In Section 3.3, it is shown how periodic excitation can be exploited in a simple algorithm for noise model identification. Finally, concluding remarks are given in Section 3.4.

3.1 Bias and Variance Effects

3.1.1 Bias Effects

Suppose the system defined by Equation 1.1 is identified within the model set described by

\[ y(t) = \hat{G}(z, \theta)u(t) + \tilde{H}(z, \theta)a(t) \]  

(3.1)

where \( \tilde{H}(z, \theta) \) is an independently parameterized noise model, which may possibly be fixed to \( H_*(z) \). The prediction errors for the model structure given in Equation 3.1 are given by:

\[ \epsilon(t, \theta) = \frac{L(z)}{\tilde{H}(z, \theta)}(y(t) - \hat{G}(z, \theta)u(t)) \]  

(3.2)

The filter \( L(z) \) is a user-selected prefilter that can be used to enhance certain frequency regions. The factor \( L(z)/\tilde{H}(z, \theta) \) always occurs in this combination, so hereinafter it will be assumed that \( L(z) = 1 \). In standard least squares prediction error identification the parameter estimates \( \theta \) are found by minimization of the objective function

\[ J_M(\theta) = \frac{1}{M} \sum_{t=0}^{M-1} \epsilon^2(t, \theta) \]  

(3.3)
Using Parseval’s relationship, Equation 3.3 may be written to provide a frequency domain description of the limiting objective function as:

\[
J_{\infty}(\theta) = \lim_{M \to \infty} \frac{1}{M} \sum_{t=0}^{M-1} E\varepsilon^2(t, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_\varepsilon(\omega) d\omega
\]  

(3.4)

where \( \Phi_\varepsilon(\omega) \) is the power spectrum of \( \varepsilon(t, \theta) \). This is a very useful expression, in that frequency domain descriptions of the bias distribution may be derived from it. For example, the bias distribution for open-loop operation using a fixed noise model \( H_*(e^{j\omega}) \), as given by Ljung (1993) is:

\[
G(z, \theta^*) = \arg \min_\theta \int_{-\pi}^{\pi} \left| G(e^{j\omega}) - \tilde{G}(e^{j\omega}, \theta) \right|^2 \frac{\Phi_u(\omega)}{|H_*(e^{j\omega})|^2} d\omega
\]  

(3.5)

In the above, \( \theta^* \) represents the limiting estimate as the number of data points tends to infinity. Equation 3.5 suggests that in open-loop conditions, the model \( \tilde{G}(e^{j\omega}, \theta) \) will approximate the true process \( G(e^{j\omega}) \) in a least squares sense, subject to the weighting factor \( \Phi_u(\omega)/|H_*(e^{j\omega})|^2 \).

An equivalent expression for the closed-loop case is also found in Ljung (1993):

\[
G(z, \theta^*) = \arg \min_\theta \int_{-\pi}^{\pi} \left| G(e^{j\omega}) - \tilde{G}(e^{j\omega}, \theta) - B(e^{j\omega}) \right|^2 \frac{\Phi_u(\omega)}{|H_*(e^{j\omega})|^2} d\omega
\]  

(3.6)

where \( B(e^{j\omega}) \) is a bias term, and the following holds:

\[
|B|^2 = \frac{\sigma^2_u \Phi_u^2(\omega)}{\Phi_u(\omega) \Phi_u(\omega)} |H(e^{j\omega}) - H_*(e^{j\omega})|^2
\]  

(3.7)

In Equation 3.7, the term \( \Phi_u^2(\omega) \) represents the power spectrum of that part of the input which is generated by the feedback, defined by Equation 1.7. Under open-loop conditions, \( \Phi_u(\omega) = 0 \), and Equation 3.6 reduces to Equation 3.5.

The presence of the bias term in Equation 3.6 is what differs closed-loop identification from its open-loop counterpart. In closed-loop, the model \( \tilde{G}(e^{j\omega}, \theta) \) will approach \( G(e^{j\omega}) - B(e^{j\omega}) \) as well as possible, subject to the weighting factor \( \Phi_u(\omega)/|H_*(e^{j\omega})|^2 \), as in the open-loop case. Therefore, the process model \( \tilde{G}(e^{j\omega}, \theta) \) will be biased unless the noise model is perfect \( (H_* = H) \). However, according to Equation 3.7 this bias effect can be made negligible, namely in frequency regions where either the feedback contribution to the input spectrum \( \Phi_u(\omega)/\Phi_u(\omega) \) is small, or where \( (\sigma^2_u/\Phi_u(\omega))^{-1} \), related to the signal to noise ratio, is high.
3.1.2 Variance Effects

Perhaps the most poorly understood aspect of closed-loop identification is the value and role of that part of the input generated from the feedback. To understand how this relates to the variance of transfer function estimates, it is necessary to examine the asymptotic distribution of parameter estimates.

Suppose the system defined by Equation 1.1 is identified within the model set described by

\[ y(t) = \tilde{G}(z, \rho)u(t) + \tilde{H}(z, \beta)a(t) \] (3.8)

and that the true system lies within this model set. Then, according to Forssell and Ljung (1999),

\[ \sqrt{M(\hat{\rho} - \rho)} \in \mathcal{N}(0, P_\rho) \] (3.9)

which means that the random vector on the left converges in distribution to the normal distribution with zero mean and covariance matrix \( P_\rho \). The asymptotic covariance matrix is given by

\[ P_\rho = \sigma^2_d (R^d_\rho + \Delta)^{-1} \] (3.10)

where \( R^d_\rho \) is related to properties of the excitation (and is equal to zero if there is no external excitation of the closed-loop system), and the term \( \Delta \) is entirely due to the noise part of the input spectrum (ie. Equation 1.7), and is always greater than or equal to zero.

While analysis of the term \( R^d_\rho \) is beyond the scope of this thesis, the \( \Delta \) term has some interesting properties which may be of consequence to experiment design:

- Since \( \Delta \geq 0 \), the noise in the loop helps to reduce the variance of the parameter estimates. The \( \Delta \) term also represents the lowest achievable accuracy, which is realized when there is no external signal present. Note however, that under these conditions, \( \Delta \) is often singular, and therefore the model parameters cannot be consistently estimated. Söderström et al. (1975) provides some special cases where \( \Delta \) is nonsingular, allowing the system to be identified using data generated from pure feedback.

- The lower limit of \( \Delta \) is zero, which occurs if the noise model order is allowed to tend towards infinity.
The upper limit of $\Delta$ is realized when the inverse of the true noise model is used as a data prefilter.

From an experiment design perspective, the part of the input signal that is generated from the noise is useful in improving the accuracy of the parameter estimates, and should therefore be accounted for. The above properties also highlight the importance of the noise model, and confirm that either a fixed (correct) noise model, or the use of an algorithm which simultaneously estimates the process and noise models are preferable in closed-loop identification. One method for estimating the process and noise models from closed-loop data is to use the Generalized Least Squares (GLS) algorithm. The idea behind the GLS method is to estimate $H(z)$ and $G(z)$ in an iterative way. In each iteration, the input-output data is first prefiltered using the noise model estimated in the previous step, followed by least squares estimation of the process model. The noise is typically modelled as an autoregressive process. The iterations are repeated until the model residuals are white, and although this convergence can be slow, the GLS algorithm is known to have favorable asymptotic properties with respect to consistency (Söderström, 1974).

### 3.2 Simulation Studies

A series of MonteCarlo simulations are used to verify some of the ideas presented above. Four sets of results are presented below, illustrating the role of external excitation and the importance of noise modelling in closed-loop identification. In each case, the true system to be identified is

$$y(t) = \frac{0.43z^{-3}}{1 - 0.92z^{-1}}u(t) + \frac{1}{1 - 0.5z^{-1}}a(t)$$

(3.11)

where $\sigma_a^2 = 1$ and a Proportional-Integral-Derivative (PID) controller is used as the feedback mechanism acting on the error $y_{sp}(t) - y(t)$. The control law is given by

$$K(z) = 0.3 \frac{1 - 1.5z^{-1} + 0.55z^{-2}}{1 - z^{-1}}$$

(3.12)

With the sampling interval set to 4 seconds and the settling time of the process equal to 284 seconds, then $N = 71$. For three of the four cases, a dither signal was injected via the setpoint, and consisted of a sum of sinusoids with equal power $\lambda_{k}^{np} = 1000$ at the FSF frequencies, $k = 0, \ldots, 6$ (see Figure 3.1). Using 1000 input-output measurements, step
response estimates were computed from an FSF model of reduced order $n = 13$. The noise model was estimated as a first order autoregressive filter:

$$
\hat{H}(z^{-1}) = \frac{1}{1 - \beta_1 z^{-1}}
$$

For each case, the same 50 different realizations of the noise source were used. In addition to the step response estimates, an average step response model was also computed. Also note that the 99% confidence bounds shown on the figures below are an average of the 50 different calculated bounds, and are plotted around the true step response of the process.

The four cases being studied are as follows:

**Case A:** Simultaneous estimation of the process and noise models; using no dither signal

**Case B:** Simultaneous estimation of the process and noise models; using a dither signal

**Case C:** Manually prefiltering the data with the incorrect noise model; using a dither signal

**Case D:** Manually prefiltering the data with the correct noise model; using a dither signal

### 3.2.1 Simultaneous Estimation of Process and Noise Models

**Case A:** The system defined by Equation 3.11 is identifiable in the absence of external excitation. Referring back to Equation 3.10, this would correspond to the special case where
$\Delta$ is nonsingular under pure feedback. It therefore follows that the identification results will represent the worst case scenario with respect to the variance of the estimates. One set of input-output data is shown in Figure 3.2. The family of 50 step response estimates obtained from data collected under pure feedback (no dither) is presented in Figure 3.3, and the corresponding average response is given in Figure 3.4. Despite there being three estimates that lie outside the confidence bounds, the estimates are still remarkably unbiased due to accurate estimation of the noise model by the FSF/GLS algorithm. In fact, the average estimate of $\beta_1$ in Equation 3.13 was 0.5132.

Case B: While the estimates obtained under pure feedback are unbiased, the variance error can be greatly reduced by the use of external excitation. A set of input-output data collected in the presence of external excitation is presented in Figure 3.5. The effect of this additional energy is illustrated in Figures 3.6 and 3.7, noting that the area enclosed by the confidence bounds is much smaller than those corresponding to pure feedback. There is also no indication of bias error, with the average estimate of $\beta_1$ being 0.5046.
Figure 3.3: Step Response Estimates Using No External Excitation, estimates (-), true response (- -), 99% confidence bounds (·)

Figure 3.4: Average Step Response Estimate Using No External Excitation, estimate (-), true response (- -), 99% confidence bounds (·)
Figure 3.5: Input-Output Data In The Presence of External Excitation

Figure 3.6: Step Response Estimates Using External Excitation, true step response (-), estimates (-), true response (---), 99% confidence bounds (--)
3.2.2 Prefiltering with an Incorrect Noise Model

Case C: An alternative to simultaneously estimating the process and noise models is to manually prefilter the input output data with the inverse of some fixed noise model, followed by direct estimation of the process model. This approach was taken to generate the results presented in Figures 3.8 and 3.9, where the effect of using an incorrect noise model is illustrated. External excitation was used, and the noise model was taken as

\[ \hat{H}(z^{-1}) = \frac{1}{1 - 0.1z^{-1}} \]  

An inherent problem with direct closed-loop identification is that the noise model must be perfect to avoid bias error in the process model. Equation 3.6 provides a frequency domain description of this bias effect. The simulation results are consistent with this theory. While the family of step response estimates presented in Figure 3.8 generally lie within the 99% confidence bounds, they are clearly not centered around the true step response estimate. In fact, a portion of the step response estimates lie outside of the 99% confidence bounds, which indicates the expected bias error. This bias error is not quite as noticeable in the
Figure 3.8: Step Response Estimates Using External Excitation and Mismatched Prefilter, estimates (-), true response (--). 99% confidence bounds (·)

averaged step response estimate, but it seems that the averaged estimate tends to be lower than the true response. This error is a direct result of prefiltering the input-output data with a mismatched noise model. However, the presence of a dither signal somewhat negates this error, namely by increasing $\Phi_u(w)$ relative to $\Phi_u^2(w)$ and $\sigma_n^2$ in Equation 3.7. The bias error would be expected to be even more significant if there was no external excitation of the closed-loop system.

3.2.3 Prefiltering with the True Noise Model

**Case D:** The use of manual prefiltering need not always result in bias error. It is not unreasonable to assume that in an industrial setting, accurate noise models have been estimated before an identification experiment. On the other hand, the task of estimating a noise model prior to an identification experiment can be made quite simple, as shown in the next section.

The results presented in Figures 3.10 and 3.11 are considered to be the best case scenario for closed-loop identification. For a given excitation signal, the most accurate process model
estimates are obtained from data that has been prefiltered by the inverse of the true noise model. Under such conditions there should be no bias error, and the variance is minimized by the combined effect of the external excitation and noise in the feedback loop.

### 3.2.4 Quantifying the Results

To complement the graphical data that has been presented within Section 3.2, two quantities that describe the accuracy of the estimates have been defined. The first, which is the average sum of squared errors of the 50 step response estimates relative to the true response, is defined as

$$\Gamma_{50} = \frac{1}{50} \sum_{m=1}^{50} \sum_{t=1}^{71} [g(t) - \tilde{g}(t, m)]^2$$  \hspace{1cm} (3.15)

where $g(t)$ is the true step response, and $\tilde{g}(t, m)$ is the step response estimated from the $m$th data set out of 50. A similar quantity is defined for the average step responses. The sum of squared errors of the average step response relative to the true response is given by

$$\Gamma_{\text{avg}} = \sum_{t=1}^{71} [g(t) - \bar{g}(t)]^2$$  \hspace{1cm} (3.16)
Figure 3.10: Step Response Estimates Using External Excitation and Correct Prefilter, estimates (-), true response (- -), 99% confidence bounds (-)

Figure 3.11: Average Step Response Estimate Using External Excitation and Correct Prefilter, estimate (-), true response (- -), 99% confidence bounds (-)
where $\bar{g}(t)$ is the average of the 50 estimated step responses. Since $\Gamma_{\text{avg}}$ is based on the averaged step response, its magnitude will largely depend on the bias error in the estimates. This is in contrast to the $\Gamma_{50}$ values, which are heavily influenced by the variance error.

The quantities defined above are presented in Table 3.1 for each of the four cases. Cases A and B correspond to simultaneous estimation of the process and noise models, with Case A using pure feedback, and Case B using external excitation. Case C corresponds to direct estimation of the process model, where an incorrect prefilter was applied to the input-output data. Lastly, Case D covers direct estimation of the process model using the correct prefilter.

Reviewing the results presented in Table 3.1, it is not surprising that Case A has the largest $\Gamma_{50}$ value. In the absence of external excitation, the variance error is large, which is also reflected in the size of the confidence envelope plotted in Figure 3.3. External excitation was applied in Cases B to D, and with the same excitation signal being used for all 3 cases, the variance error can be considered a constant. Of these 3 cases, Case C has the largest $\Gamma_{\text{avg}}$ value, which can be attributed to the bias error resulting from prefilter mismatch.

Also notable in Table 3.1 is that Case D outperformed Case B. Section 3.2.3 asserts that prefiltering input-output data with the inverse of the true noise model is theoretically the best case scenario, and the simulation results seem to support this. It was expected that the limiting $\Gamma$ values would correspond to those calculated for Case D, with the performance of Case B approaching, but not exceeding that of Case D. The difference between the $\Gamma$ values for these two cases is relatively small, proving that simultaneous estimation of the process and noise models can be an equally effective approach to closed-loop identification.

### 3.3 The Use of Averaging to Estimate the Noise Model

The use of periodic excitation is most commonly associated with frequency domain identification. However, when used in time domain identification, it offers several advantages, as noted by McKelvey (1996). When periodic excitation is applied to linear closed-loop
systems, all signals in the loop become periodic after a transient period. This property forms the basis for a simple averaging technique in which the data set may be reduced into a single period. One of the unique advantages of this is that the noise properties may be recovered from the input-output data. This allows for independent estimation of a noise model, which may subsequently be used as a prefilter in a closed-loop identification routine for obtaining an estimate of the process model.

The method proposed by McKelvey (1996) only requires that the noise has a sample mean equal to zero, and cannot be a periodic signal with a period time equal to a fraction of the period time of the excitation signal. With reference to the system in Figure 1.1, this would imply that the sequence \( a(t) \) must be zero mean white noise (a standing assumption), and also that the noise filter \( H(z) \) must be stable. However, the proposed method may be extended to include non-stationary noise sequences by appropriately differencing the input-output data. In addition, there is no restriction on the periodic waveform, so such an experiment could be performed easily with the use of a single sine wave injected at the setpoint or added to the controller output. Given this, consider the situation in which a set of noisy measurements have been taken from one of the signals in the feedback loop. The following then holds:

\[
x(t) = \hat{x}(t) + w(t)
\]  
(3.17)

In the above, \( \hat{x}(t) \) is the noise-free and periodic component of the signal, and \( w(t) \) is an unknown noise source which acts additively on the measured variable, \( x(t) \).

Assume that an excitation signal with a period of \( P \) has been applied to the system, and that \( M \) measurements have been taken (\( M \) being selected such that an integer number of periods, \( M_P \), have been measured). The data set may be reduced into a single period of averaged data, as in

\[
x_P(t) = \frac{1}{M_P} \sum_{k=0}^{M_P-1} x(t + kP)
\]  
(3.18)

for \( t = 0, \ldots, P - 1 \). If the assumption that the noise source \( w(t) \) is a zero mean random variable holds, then it follows that

\[
\lim_{M_P \to \infty} x_P(t) = \hat{x}(t)
\]  
(3.19)

for \( t = 0, \ldots, P - 1 \). That is, \( x_P(t) \) is a consistent estimate of the noise-free component \( \hat{x}(t) \).
With a consistent estimate of \( \bar{x}(t) \) available, it is possible to recover the noise signal \( w(t) \). It is clear from Equations 3.17 and 3.19 that

\[
\bar{w}(t) = x(t) - x_P(t)
\]  

where \( \bar{w}(t) \) is a consistent estimate of \( w(t) \), and \( x_P(t) \) is periodically continued outside of \( t = 0, \ldots, P - 1 \). With this estimate of the noise component, time series analysis may then be used to calculate a noise model.

When applied to the process output of a closed-loop system, the terms in Equation 3.17 become

\[
x(t) = y(t)
\]

\[
\ddot{x}(t) = G(z)K(z)S(z)y_{xp}(t)
\]

\[
w(t) = v(t) - G(z)K(z)S(z)v(t)
\]

In Equation 3.23, the term \( G(z)K(z)S(z)v(t) \) represents the effect of the noise \( v(t) \) passing through the feedback loop. Note that Equation 3.23 becomes

\[
w(t) = v(t)
\]

\[
= H(z)a(t)
\]

subject to the assumption that

\[
v(t) \gg G(z)K(z)S(z)v(t)
\]

Equation 3.24 suggests that, under this assumption, the non-periodic element in the measured output is generated by \( H(z)a(t) \), permitting estimates of \( H(z) \). An estimate of the noise variance may then be computed from

\[
\sigma_a^2 = \text{var}[H^{-1}(z)w(t)]
\]

Two sets of MonteCarlo simulations are used to examine these ideas. The first simulation study is concerned with identifying the noise model from closed-loop data collected from the same system described in Section 3.2. In the second simulation study, the noise model is changed to

\[
H(z) = \frac{1}{(1 - z^{-1})(1 - 0.2z^{-1})}
\]
In both simulations, a sinusoid with an amplitude of 2 at the 3rd FSF frequency \( \omega_3 = \frac{\pi}{4} \) rad/sample) is applied as the reference signal. The data length is selected as 983 samples, corresponding to 39 periods. Prior to using the algorithm described above for noise model estimation, the first \( N = 71 \) data points are discarded, eliminating any transients in the data set. Additionally, the stationary part of the noise model is identified within the class of first order autoregressive filters of the form

\[
\hat{H}(z^{-1}) = \frac{1}{1 - \beta_1 z^{-1}}
\]  

(3.28)

where the parameter \( \beta_1 \) is estimated using a least squares approach. For both sets of simulations, 250 realizations of the noise \( a(t) \) are used.

Presented in Figure 3.12 is a set of data collected in the first set of simulations. The average estimate of \( \beta_1 \) was 0.5259, which is very close to the true value of 0.5000. In addition, for each of the 250 estimated noise models, the noise spectrum \( |H(e^{j\omega_k})|^2 \) at the frequencies \( \omega_k \) for \( k = 0, \ldots, 6 \) was calculated. The average estimated noise spectrum is compared to the true noise spectrum in Table 3.2 and Figure 3.13. The accuracy of the estimated noise model and spectrum shows that the use of periodic excitation coupled with the averaging routine described above is an effective approach to identifying the noise properties of closed-loop systems.

<table>
<thead>
<tr>
<th>( k )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>H(e^{j\omega_k})</td>
<td>^2 )</td>
<td>4.000</td>
<td>3.938</td>
<td>3.765</td>
<td>3.508</td>
<td>3.205</td>
</tr>
<tr>
<td>(</td>
<td>\hat{H}(e^{j\omega_k})</td>
<td>^2 )</td>
<td>4.495</td>
<td>4.411</td>
<td>4.179</td>
<td>3.844</td>
<td>3.460</td>
</tr>
</tbody>
</table>

Table 3.2: Average Estimated Noise Spectrum: Stationary Noise Case

To validate the assumption given by Equation 3.25, a sample of the signals \( v(t) \) and \( v(t) - G(z)K(z)S(z)v(t) \) are compared in Figure 3.14. The signals are similar, suggesting that the contribution of \( G(z)K(z)S(z)v(t) \) to the noise in the output is small. It is also important to note that the variance of \( v(t) \) is estimated to be 1.34, while the variance of \( G(z)K(z)S(z)v(t) \) is 0.14.

To illustrate the application of the averaging routine on systems affected by a non-stationary disturbance, the MonteCarlo simulations were repeated for the case where the noise model is given by Equation 3.27. The noise filter given by Equation 3.27 has a pole on the unit circle, meaning that the disturbance is non-stationary in the mean. Figure 3.15
Figure 3.12: Sample Excitation and Output Signals: Stationary Noise Case

Figure 3.13: Average Noise Spectrum: Stationary Noise Case, true (○), actual (*)
Figure 3.14: Sample Noise Signals: Stationary Noise Case, $v(t)$ (-), $v(t) - G(z) K(z) S(z) v(t)$ (-)

presents a sample of data collected in this set of simulations. Prior to applying the averaging routine, the output data is first differenced. The averaging routine is then used to estimate the stationary component of the noise:

$$H(z^{-1}) = \frac{1}{1 - 0.2z^{-1}}$$

(3.29)

The average of 250 estimates of $\beta_1$ was determined to be 0.2383, which is comparable to the true value of 0.2000. The average noise spectrum was also calculated, and is presented in Table 3.3 and Figure 3.16. Although the estimate of the noise spectrum is not as accurate as that for the stationary noise case (Figure 3.13), the results are satisfactory.

<table>
<thead>
<tr>
<th>$k$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>H(e^{j\omega_k})</td>
<td>^2$</td>
<td>1.562</td>
<td>1.559</td>
<td>1.547</td>
<td>1.529</td>
<td>1.504</td>
</tr>
<tr>
<td>$</td>
<td>\tilde{H}(e^{j\omega_k})</td>
<td>^2$</td>
<td>1.734</td>
<td>1.728</td>
<td>1.711</td>
<td>1.683</td>
<td>1.647</td>
</tr>
</tbody>
</table>

Table 3.3: Average Estimated Noise Spectrum: Non-Stationary Noise Case
Figure 3.15: Sample Excitation and Output Signals: Non-Stationary Noise Case

Figure 3.16: Average Noise Spectrum: Non-Stationary Noise Case, true (o), actual (*)
3.4 Summary

This chapter highlighted the role of the noise model in direct closed-loop identification. While there has been much emphasis on the topic in the literature, it is still often misunderstood from a practical standpoint.

In Section 3.1, theoretical expressions for bias and variance errors in estimates obtained from prediction error methods were presented. It was shown that unbiased estimates of the process model can be computed from open-loop data, regardless of the noise model. This is in contrast to the closed-loop case, where it was shown that erroneous noise modelling results in bias of the process model estimate. This effect was verified by simulation studies, and is most notable in the results of Section 3.2.2.

The relation between the noise model and the variance of parameter estimates was also covered in this chapter. It was suggested that a fixed (and correct) noise model should maximize the value of that part of the input originating from the noise. In other words, the presence of the feedback mechanism helps in reducing the variance of the identified model.

The only situation in which this is not true is when the noise model order is allowed to tend towards infinity.

In Section 3.3, a simple algorithm for noise model estimation was reviewed. These ideas were originally presented by McKelvey (1996), who advocated the use of periodic excitation in closed-loop experiments as an efficient means of reducing variance effects. The estimated noise model could then be used as a prefilter in future identification experiments. A series of MonteCarlo simulations were used to illustrate the application of the algorithm on systems affected by stationary or non-stationary disturbances.
Chapter 4

The Two-Phase Approach to Optimal Closed-Loop Experiments

This chapter addresses the problem of optimal experiment design for closed-loop identification experiments. Typically, the objective of an optimal identification experiment is to minimize some function of the variance in the estimated process model. The focus of this chapter will be a design methodology which attempts to minimize the variance in FSF-generated step response estimates obtained from closed-loop data. It is shown that the optimal experiment design is dependent on unknown parameters of the closed-loop system, eliminating the possibility of off-line design prior to an identification experiment. The proposed two-phase approach to optimal closed-loop identification, presented in this chapter, utilizes periodic excitation as a means of efficiently estimating the parameters of the closed-loop system that are necessary for a feasible experiment design. By making a simple adjustment to the excitation signal at some intermediate point during the experiment, the objective of optimal experiment design can be satisfied.

4.1 Optimal Test Signal Design

Consider the problem of obtaining the optimal FSF spectrum ratios, given in Table 2.1, in a single closed-loop experiment. The FSF spectrum reflects the energy of the input signal at the FSF frequencies. Under closed-loop conditions, the input may not be directly adjusted, as it is manipulated by a feedback mechanism. However, the spectral properties of the input signal $u(t)$ can be influenced by manipulating $y_p(t)$ in an appropriate fashion. Recall
Equations 1.3 and 1.5:

\[ u(t) = S(z)K(z)y_{sp}(t) - H(z)S(z)K(z)a(t) \]  
(4.1)

\[ \Phi_u(\omega) = |S(e^{j\omega})|^2 |K(e^{j\omega})|^2 \Phi_{y_{sp}}(\omega) + |S(e^{j\omega})|^2 |K(e^{j\omega})|^2 |H(e^{j\omega})|^2 \sigma_o^2 \]  
(4.2)

An analogous expression can be written to describe the FSF spectrum \( \lambda_k \) obtained from a closed-loop experiment as follows:

\[ \lambda_k \approx |S(e^{j\nu_k})|^2 |K(e^{j\nu_k})|^2 \lambda_{k}^{y_{sp}} + \lambda_{k}^{fb} \]  
(4.3)

for \( k = 0, \ldots, \frac{n-1}{2} \). See Appendix B for details on the derivation of Equation 4.3. The term \( \lambda_{k}^{y_{sp}} \) represents the FSF spectrum of the reference signal \( y_{sp}(t) \):

\[ \lambda_{k}^{y_{sp}} = \sum_{t=0}^{M-1} \phi_{y_{sp}}^k(t)\phi_{y_{sp}}^k(t) \]  
(4.4)

where

\[ \phi_{y_{sp}}^k(t) = \frac{1}{N} \frac{1 - z^{-N}}{1 - e^{j\nu_k}z^{-1}} y_{sp}(t) \]  
(4.5)

The term \( \lambda_{k}^{fb} \) in Equation 4.3 is the feedback contribution to the FSF spectrum. While Equation 4.3 is only an approximation (for reasons explained later in Section 4.5.2), it will be shown to be an effective guide for shaping \( \lambda_{k}^{y_{sp}} \) to achieve a desired FSF spectrum and is the cornerstone for the work in this chapter.

The transfer function from the setpoint to the input signal in Equation 4.1, which will be denoted by \( Q(z) \), is a function of the unknown process, \( G(z) \).

\[ Q(z) = S(z)K(z) \]  
(4.6)

\[ = \frac{K(z)}{1 + G(z)K(z)} \]  
(4.7)

Therefore, in order to shape the input spectrum using the setpoint as the excitation signal, the amplitude response of \( Q(z) \) will need to be estimated. However, according to Equation 4.3, it is only the relatively small number of weights \( |Q(e^{j\nu_k})|^2 \) for \( k = 0, \ldots, \frac{n-1}{2} \), that need to be known. This simplification arises because the FSF model structure being used to represent the process \( G(z) \) is a set of parallel bandpass filters, with each filter selecting a particular frequency component from the process input, namely \( \omega_k \) for \( k = 0, \ldots, \frac{n-1}{2} \). It should also be noted that in the context of the FSF model structure, a persistently exciting
excitation signal is one which has energy at the \( n \) FSF frequencies. Persistent excitation is a necessary condition for consistent estimation of the process model\(^1\), and while there may be excitation over the full range \([-\pi, \pi]\), any energy away from the centre frequencies of the FSF model is attenuated.

Chapter 3 highlighted the role of the noise model with respect to bias and variance effects on the estimated process model. It was suggested that having prior knowledge of the noise model is advantageous, but not necessary, since the use of a GLS algorithm can efficiently estimate both the process and noise models. However, the design of an experiment for optimal FSF model identification, whether it be in open-loop or closed-loop, cannot be done without prior knowledge of the noise model. The optimal FSF spectrum ratios given in Table 2.1 define a desired energy distribution in the prefiltered input. The effect of data prefiltering on the FSF spectrum energies \( \lambda_k \) can be described as follows:

\[
\lambda'_k = \frac{\lambda_k}{|H(e^{j\omega})|^2}
\]

(4.8)

where \( \lambda'_k \) for \( k = 0, \ldots, \frac{n-1}{2} \) represents the \( k \)th FSF spectrum energy of the input variable \( u(t) \) after prefiltering with the inverse of the noise model. By rearrangement of Equation 4.8 and replacing \( \lambda'_k \) with a set of desired energies \( \lambda_k^{\text{opt}} \) having the distribution defined in Table 2.1, it follows that an input signal with a FSF spectrum of

\[
\lambda_k = \left|H(e^{j\omega_k})\right|^2 \lambda_k^{\text{opt}}
\]

(4.9)

would be considered optimal.

The presence of noise in the feedback loop must also be accounted for when designing an experiment for optimal FSF model identification. The spectrum of this stochastic part of the input is given by Equation 1.7, or in terms of the FSF model structure, it is described by \( \lambda_k^{fb} \) in Equation 4.3. The set of \( \lambda_k^{fb} \) for \( k = 0, \ldots, \frac{n-1}{2} \) will be referred to here as the 'background spectrum'. This background spectrum is considered to be beneficial in direct closed-loop identification, helping to reduce the variance of the parameter estimates, as discussed in Section 3.1.2.

Suppose that the three items listed above are known. That is, assume that the amplitude response of \( Q(z) \), the noise model \( H(z) \), and the background spectrum are given. In

\(^1\)Technically, persistent excitation ensures that the matrix \((\Phi^*\Phi)^{-1}\) exists. See Equation 2.18.
addition, assume that the settling time $N$ of the open-loop process $G(z)$ is known. The optimal FSF spectrum of the reference signal is then given by:

$$
\lambda_k^{\text{opt}} = \frac{|H(e^{j\omega_k})|^2 \lambda_k^{\text{opt}} - \lambda_k^b}{|Q(e^{j\omega_k})|^2} \tag{4.10}
$$

Equation 4.10 can be verified by substituting for $\lambda_k^{\text{opt}}$ in Equation 4.3, which can then be simplified to Equation 4.9.

Subject to prior knowledge of $N$, $|Q(e^{j\omega_k})|^2$, $|H(e^{j\omega_k})|^2$, and the background spectrum, optimal test signal design for FSF model identification is straightforward. Upon calculating the required reference spectrum from Equation 4.10, the procedure for multisine design described in Section 2.6 may be used to realize a reference signal with the necessary spectral distribution. However, prior knowledge of the closed-loop system is generally not available to this extent. While a separate experiment could be used to estimate the unknown parameters in Equation 4.10, the cost of such an approach makes this unattractive.

An alternative approach proposed here is to estimate $|Q(e^{j\omega_k})|^2$, $|H(e^{j\omega_k})|^2$, and the background spectrum at some intermediate point during the optimal experiment. That is, a single experiment can be divided into two phases, each with different objectives. Under this proposed scheme, the first phase of the experiment is concerned with identifying the amplitude response of $Q(z)$ and the noise model, as well as measuring the feedback contribution to the input spectrum. In other words, the first phase is used in replacement of a separate identification experiment. In the second phase of the experiment, any additional energy required to achieve the optimal FSF spectrum may then be added, using Equation 4.10 to determine the necessary spectral properties of the reference signal. That is, the input-output data of both phases of the experiment are used to estimate the final process model $G(z)$. It is this characteristic that distinguishes the two phase approach from the idea of using a separate experiment to estimate the unknown parameters in Equation 4.10.

The timeline of the proposed two-phase approach is outlined in Figure 4.1. The length of the first phase of the experiment is denoted by $M_1$, and the length of the complete experiment is indicated by $M$. The length of the second phase of the experiment will be denoted by $M_2$ (not shown in the figure), and is equal to $M - M_1$. The objective of the experiment is to obtain a desired set of FSF energies $\lambda_k = \lambda_k^{\text{opt}}$ for $k = 0, \ldots, n-1$ at time $t = M$. 
4.2 Phase One

The objective of the first phase of the experiment is to estimate the amplitude response of the weighting function $Q(z)$, the noise properties, and the background spectrum as efficiently as possible. The input energy in this phase of the experiment will also form part of that used to estimate the process model $\hat{G}(z)$ upon completion of the two-phase experiment. It is therefore critical that the first phase of the experiment is designed carefully. The data collected in the first phase of the experiment must include valuable information on $Q(z)$, $H(z)$, and $\lambda_k^b$, as well as the process model $G(z)$, at the frequencies of interest. The remainder of this section focuses on how periodic excitation can be used to meet the objectives of the first phase of the experiment. It will be shown that a periodic reference signal composed of a sum of sinusoids at the FSF frequencies $\omega_k$ for $k = 0, \ldots, \frac{n-1}{2}$ will provide sufficient information for the estimation of $|Q(e^{j\omega_k})|^2$, $H(z)$, and $\lambda_k^b$ at the end of the first phase. In fact, the use of periodic excitation will allow for a particular set of estimation routines that when considered as a whole, form a highly efficient set of computations. Furthermore, with the energy concentrated in the specified FSF frequencies, the set of $M$ data points collected from the complete experiment will be very useful in the estimation of the process model.

4.2.1 Estimation of the Weighting Function

One objective in the first phase of the optimal experiment is to estimate the amplitude response of the weighting function $Q(z)$. The process input $u(t)$ may be expressed in terms of the weighting function by substituting Equation 4.6 into Equation 4.1, which gives

$$u(t) = Q(z)y_{sp}(t) - H(z)S(z)K(z)a(t)$$  \hspace{1cm} (4.11)
Since \( y_{tp}(t) \) is uncorrelated with the noise term, the weighting function \( Q(z) \) may be identified in an open-loop way, using measurements of \( y_{tp}(t) \) and \( u(t) \) generated by the system defined in Equation 4.11. This is similar to the first step taken in the joint input-output identification routines described in Section 1.3.1. However, identification of the entire transfer function \( Q(z) \) is not necessary in the proposed two-phase approach to optimal experiments; only the amplitude response \( |Q(e^{j\omega_k})|^2 \) for \( k = 0, \ldots, \frac{n-1}{2} \) is needed. Therefore, it seems appropriate to use the FSF model structure to represent \( Q(z) \) since the parameters of this particular model structure are equal to the process frequency response, \( Q(e^{j\omega_k}) \).

The use of the FSF model structure to estimate frequency response coefficients at arbitrary frequencies differs slightly from its conventional use as a process model. When used as a process model, the parameter \( N \) is selected to be equal to the process settling time, as described in Section 2.1. However, in situations where it is not necessary to completely describe the process dynamics, \( N \) can be chosen such that the frequency response of the unknown system is estimated at the frequencies \( \frac{2\pi k}{N} \) for \( k = 0, \ldots, \frac{n-1}{2} \). So, when using the FSF model to estimate \( Q(e^{j\omega_k}) \), \( N \) may be set equal to the settling time of the open-loop process \( G(z) \). Choosing \( N \) on this basis will ensure that \( Q(e^{j\omega_k}) \) is identified at the frequencies corresponding to the specified \( \lambda_k^{opt} \) in Equation 4.10. It has already been assumed that an estimate of the process settling time \( N \) is available.

With the identification of \( Q(z) \) representing an open-loop problem, it follows from Ljung (1999) that under certain conditions, consistent estimates of \( Q(e^{j\omega_k}) \) may be obtained irrespective of the noise contribution. This holds true when the process and noise models are independently parameterized, as in

\[
 u(t) = \tilde{Q}(z, \theta)y_{tp}(t) - \tilde{H}_u(z, \rho)\alpha(t) \tag{4.12}
\]

where \( \tilde{H}_u(z, \rho) \) is a noise model with parameter vector \( \rho \). The two-phase approach may incorporate a GLS algorithm to estimate \( Q(e^{j\omega_k}) \) along with the associated noise model in Equation 4.1. The effect of modelling the noise \( \tilde{H}_u(z, \rho) \) in the two-phase approach is investigated by simulation studies presented in Section 4.4.3.

The outlined approach to the estimation of \( Q(e^{j\omega_k}) \) considerably simplifies the design of the reference signal to be used in the first phase of the experiment. Although \( Q(e^{j\omega_k}) \) and \( G(e^{j\omega_k}) \) will be identified in different phases of the experiment, they will be estimated at
the same set of frequencies, namely \( \frac{2\pi k T}{N} \) for \( k = 0, \ldots, \frac{N-1}{T} \). Therefore, the reference signal for the entire experiment only needs to contain energy at these frequencies.

### 4.2.2 Estimation of the Noise Model

Knowledge of the noise properties is essential for the design of optimal closed-loop experiments. The optimal dither spectrum given by Forssell and Ljung (2000), shown in Equation 1.23, is a function of the noise spectrum, \( \Phi_n(\omega) \). The optimal reference spectrum for FSF model identification is given in Equation 4.10, where it is shown to depend on \( |H(e^{j\omega})|^2 \). It is therefore necessary to estimate the noise model in the first phase of the experiment.

Since the excitation signal used in the first phase of the experiment is periodic, the measured output \( y(t) \) will also be periodic after a transient period. The averaging routine described in Section 3.3 will be used to reduce the set of output measurements \( y(t) \) for \( t = 0, \ldots, M_1 - 1 \) into a single period of data, as in Equation 3.18. A consistent estimate of the additive noise disturbance \( H(z)a(t) \) can then be computed from Equation 3.20. The noise will be modelled using an autoregressive model structure, with the model order being determined by the PRESS. That is,

\[
\hat{\Phi}(z) = \frac{1}{1 - \beta_1 z^{-1} - \beta_2 z^{-2} - \ldots}
\]  

Identification of the noise model in the first phase of the experiment represents the second use of measurements collected from \( t = 0, \ldots, M_1 - 1 \). In Section 4.2.1, it was the specified energy distribution of the reference signal that allowed for the estimation of \( Q(e^{j\omega}) \), whereas in this section it is simply the periodicity of the reference signal that allows the averaging routine to be used.

### 4.2.3 Estimation of the Background Spectrum

Under closed-loop conditions, the process input is formed by the sum of a deterministic component due to the setpoint and a stochastic component due to the noise, as shown in Equation 4.1. The energy contribution of the stochastic component is considered to be valuable, as it helps to reduce the variance of the parameter estimates. It is therefore necessary to somehow measure the contribution this stochastic element of the input makes to the FSF spectrum. With the reference signal in the first phase of the experiment being periodic, averaging of the process input \( u(t) \) will be used to recover the contribution of the
noise to the signal. Measurement of the background spectrum from the first phase of the experiment will allow for a more informed prediction of $\lambda^{fb}_k$, the background spectrum for the whole experiment.

Consider the situation where a set of input measurements have been collected from a closed-loop experiment where the reference signal is periodic. From Equation 4.1, it is clear that the input measurements are corrupted by a noise source. That is,

$$u(t) = \tilde{u}(t) + w_u(t)$$  \hfill (4.14)

where $\tilde{u}(t)$ is the noise-free and periodic component of the input signal, and $w_u(t)$ is an unknown noise source which acts additively on the measured input, $u(t)$. Assuming that the noise source $w_u(t)$ is a zero mean random variable, a consistent estimate of $\tilde{u}(t)$ can be made by reducing the data set into a single period of averaged data, analogous to Equation 3.18. The noise sequence $w_u(t)$ can then be estimated in a manner similar to that given in Equation 3.20.

Comparing Equations 4.1 and 4.14, it is clear that

$$\tilde{u}(t) = S(z)K(z)y_{sp}(t)$$  \hfill (4.15)

and

$$w_u(t) = -H(z)S(z)K(z)a(t)$$  \hfill (4.16)

That is, the periodic portion of the measured input is simply the realization of passing the reference signal $y_{sp}(t)$ through a stable linear filter, $S(z)K(z)$. Similarly, the noise in the input measurements is generated by $a(t)$ passing through the filter $H(z)S(z)K(z)$. However, time series analysis will not be applied here to the estimate of $w_u(t)$ in an attempt to model $H(z)S(z)K(z)$, which is in contrast to the application of the averaging routine in the estimation of $H(z)$. Instead, the FSF energy contained in the estimated $w_u(t)$ will be calculated at the frequencies $\frac{2\pi k}{N}$ for $k = 0, \ldots, \frac{N-1}{2}$, with the parameter $N$ being selected as the settling time of the process $G(z)$. That is,

$$\bar{\lambda}^{fb}_k = \sum_{i=0}^{M_1-1} \phi^{fb*}_{01}(t) \phi^{fb}_{01}(t)$$  \hfill (4.17)

where $M_1$ is the length of the first phase of the experiment, and

$$\phi^{fb}_{01}(t) = \frac{1}{N} \frac{1 - z^{-N}}{1 - \rho w_a z^{-1}} \tilde{u}(t)$$  \hfill (4.18)
where \( \hat{\omega}_u(t) \) is an estimate of the noise \( \omega_u(t) \), as calculated from the averaging routine.

The calculated energy \( \hat{\lambda}_k^{fb1} \) represents the feedback contribution of the noise to the input spectrum in the first phase of the experiment. Since the FSF spectrum energies are proportional to the number of data points (see Equation 2.30), the contribution of the feedback to the input spectrum for the whole experiment can be estimated as

\[
\hat{\lambda}_k^{fb} = \frac{M}{M_1} \hat{\lambda}_k^{fb1}
\]

where \( M \) is the total length of the two phase experiment. Alternatively, the feedback contribution in the second phase of the experiment can be predicted via

\[
\hat{\lambda}_k^{fb2} = \frac{M_2}{M_1} \hat{\lambda}_k^{fb1}
\]

where \( M_2 \) is the length of the second phase of the experiment.

4.2.4 Design of Phase 1 Excitation

The objective of the first phase of the experiment is to estimate the amplitude response of the weighting function, the noise model, and the background spectrum as accurately as possible. The quality of the estimates made in the first phase strongly affects the ability to optimally design the experiment as a whole, and as such, the excitation signal used in this phase of the experiment should be carefully designed. Estimation of the amplitude response of \( Q(z) \), as outlined in Section 4.2.1, requires that the reference signal excites at least the FSF frequencies \( \omega_k \) for \( k = 0, \ldots, \frac{n-1}{2} \). Energy at non-FSF frequencies is deemed to be wasteful, namely because the narrow bandpass nature of the FSF model structure attenuates energy away from the centre frequencies of the FSF filters. The averaging routine described in Section 3.3 is utilized to estimate the noise model and the background spectrum. Therefore, the reference signal must be periodic. Combining the requirements of the framework outlined in Sections 4.2.1 to 4.2.3, a suitable reference signal for the first phase of the experiment is one which is periodic and excites only the FSF frequencies.

Within the set of periodic signals that excite a relatively few number of frequencies, there remains a significant amount of flexibility for design. However, careful consideration must be given to possible constraints on the total energy of the reference signal, as well as the spectral distribution of this energy among the FSF frequencies. In addition, the period
length and the total length of the first phase need to be chosen with care. The purpose of this section is to provide some insightful discussion on the design of phase one excitation, as this is the lesser 'automated' phase of the two.

To introduce some notation, first recall that the length of the first phase is denoted as $M_1$. The reference signal will be composed of an integer number of periods, $M_{P_1}$, with the period length denoted as $P_1$. Recalling Equations 3.18 and 3.19, it is reasonable to suggest that the accuracy of the averaging routine should improve as $M_{P_1}$ increases. With respect to the first phase of the experiment, this means that the noise model and background spectrum can be estimated more reliably. Thus, the period length $P_1$ should be made as small as possible, such that the signal can be composed of the greatest number of periods. When using the van den Bos algorithm described in Section 2.6 for multisine signal design, it is recommended that the period length be equal to an integer multiple of the process settling time (this ensures that $r$ in Equation A.8 is an integer). Therefore, it would seem that the best choice for $P_1$ is the process settling time $N$.

The specification of a desired length for the first phase of the optimal experiment represents a trade-off. If a larger portion of the total experiment is allotted to the first phase, improved estimates of the noise model, weighting function, and background spectrum will be obtained. The drawback in this situation is that the energy required in the second phase of the experiment will have to be injected into the system over a relatively small number of samples. Should there be a constraint on the output variance, or some upper bound on the output, it would most likely be exceeded if the duration of the second phase is small. To be conservative, a reasonable choice for $M_1$ would be half of the total experiment length $M$. An additional benefit of setting $M_1$ to be equal to half of the total experiment length is that the prediction of the total background spectrum $\lambda_k^{fb}$ is simplified. Specifically, the total background spectrum $\lambda_k^{fb}$ can be estimated by doubling the calculated phase one background spectrum, $\hat{\lambda}_k^{fb}$.

The final consideration that must be made with respect to the first phase excitation is the spectral distribution of its energy among the FSF frequencies. Referring to the four step procedure for multisine signal design given in Section 2.6, it will be necessary to specify a desired FSF spectrum for the reference signal. Careful consideration must not only be given to the relative magnitudes of the desired FSF energies, but also the total energy. Naturally,
It is desirable to inject as much energy as possible in this phase of the experiment in order to reduce the variance of the estimated weighting function $\hat{Q}(e^{j\omega_k})$. The maximum energy that may be injected will be limited by any constraints on the range that the reference signal (and thus the output) can span. For example, suppose there is a constraint $A_{\text{max}}$ on the amplitude of the reference signal. Substituting $A_{\text{max}}$ for $A_k$ in Equation 2.36 gives

$$\lambda_{k,\text{max}} = \frac{A_{\text{max}}^2 M_1}{4}$$

(4.21)

where $\lambda_{k,\text{max}}$ is the maximum energy at any FSF frequency of the reference signal such that the constraint on the amplitude is satisfied. It should be noted that Equation 4.21 will hold true only when the reference signal is composed of a single sine wave. Excitation of more than one frequency requires the use of a multisine reference signal. While the peak amplitude of a multisine depends on the amplitude and phase shift of the individual sine waves, it is reasonable to say that in most cases it will exceed the maximum amplitude of the individual sine waves. Therefore, Equation 4.21 may only be used as a guideline in the specification of the reference signal spectrum. In addition, since the van den Bos algorithm for multisine signal design has no facility for the definition of a constraint on the peak amplitude, a trial-and-error approach may need to be used when applying it to the reference signal in this phase of the experiment.

The most suitable shape for the reference signal's FSF spectrum is debatable. The following two arguments should be considered:

- in the first phase of the experiment, it is only the frequency response of $Q(z)$ that is being estimated, and not its step response. Thus, it is not necessary to use a test signal that has the optimal FSF spectrum shape described in Section 2.5. Instead, the estimates of $Q(e^{j\omega_k})$ at all FSF frequencies will be considered to be equally important, and therefore the reference spectrum should be flat.

- The overall objective of the experiment is for the filtered input signal to have the optimal FSF spectrum shape. Since the optimal FSF spectrum shape essentially defines a scale of importance for the parameter estimates, it would seem reasonable to apply the same scale to the estimation of $Q(e^{j\omega_k})$ in the first phase of the experiment. For example, the results of Section 2.5 identify the energy at the zero frequency $\omega_0$
to have the strongest influence on the variance of the estimated step response (of the process \( G(z) \)). It therefore seems critical that a desired value of \( \lambda^\text{opt}_0 \) is actually achieved. From Equation 4.10, it is clear that the ability to inject precisely the desired amount of energy \( (\lambda_0^* = \lambda^\text{opt}_0) \) will depend on the accuracy of the estimate \( Q(e^{t\omega_0}) \). On the other hand, energy at the 6th FSF frequency \( \omega_6 \) has the least effect on the variance of the step response estimate, and therefore it is not necessary to be as precise with the energy content of the reference signal at this frequency as it is for the zero frequency.

While both arguments offer some valid reasoning, it is the second argument that will be supported in this thesis. The use of a flat spectrum in the first phase of the experiment would present difficulties in the design of the second phase. In particular, it is quite likely that the reference signal in the second phase would have to contain a significantly larger proportion of energy in the low frequencies, with little or none in the high frequency region. Any excitation signal with such a skewed energy distribution is likely to violate the constraints on the peak amplitude.

The complication with the second approach is that the noise in the system defined by Equation 4.1 is not white, and the noise model \( H(z)S(z)K(z) \) is unknown. Therefore it will be impossible to design a reference signal that will result in optimal estimation of the weighting function after prefiltering has been applied. While the reference signal will not have the optimal spectrum shape, it will still be designed to contain a larger portion of its energy in the lower frequencies.

### 4.3 Phase Two

The objective of the second phase of the experiment is to ensure that the input signal prefiltered by the inverse of the noise model, when considered over the entire experiment, is optimal in the sense defined in Section 2.5. While the first phase of the experiment is used primarily for the estimation of the weighting function, noise model, and background spectrum, the open-loop process \( G(z) \) is also being excited at the FSF frequencies \( \omega_k \) for \( k = 0, \ldots, \frac{n-1}{2} \). In other words, the data set \( \{u(t), y(t)\} \) for \( t = 0, \ldots, M_1 - 1 \) contains valuable information about \( G(z) \). Therefore, the data from the first phase of the experiment
will not be discarded, but instead will be supplemented with the data collected in the second phase. In fact, using knowledge gained from the first phase of the experiment, the desired spectral properties of the excitation to be used in the second phase can be defined rather precisely.

The second phase of the experiment begins immediately after the estimates computed from the first phase become available. With the substantial computing power of modern microprocessors, the calculation of $\hat{Q}(e^{j\omega_k})$, $\hat{H}(z)$, and $\lambda_k^{opt}$ can be performed in only a few seconds. Therefore, the proposed two phase methodology to optimal closed-loop identification can be regarded as an on-line approach. In situations where the sampling interval is relatively large, the calculations performed between the two phases can be made without interruption of the experiment. However, if the calculations take longer than one sampling interval, then external excitation of the closed-loop system will be absent until the computed reference signal for phase two becomes available. Setting the reference signal to zero or holding it at its present value during this intermediate step will not adversely affect the quality of the data set in any appreciable way.

4.3.1 Fixing the Total Input Energy

The optimal FSF spectrum ratios given in Table 2.1 define the distribution of energy among the FSF frequencies that will minimize the variance error in the estimated step response for a fixed level of input energy. The area enclosed by the confidence bounds around the step response estimate is affected by the total energy in the input signal. Naturally, a higher level of input energy will reduce the variance of the estimate. The total energy in the input can be fixed by specifying any one of the desired FSF energies $\lambda_k^{opt}$, with the remaining energies being set in accordance with the optimal distribution.

The selection of total energy was also an issue considered in the first phase of the experiment. It was suggested that a trial-and-error application of the van den Bos algorithm could be used to find a test signal that contains the maximum amount of energy without violating any constraints on the peak amplitude. A similar approach cannot be used in the second phase of the experiment, as it would be far too time consuming. That is, there would be an extended period of time (more than a few sampling intervals) between the two phases where there would be no external excitation.
Instead, the energy desired at the zero frequency \( \lambda_0 \) over the entire experiment will be fixed by allowing the user to specify a tolerable error in the process gain estimate, \( \hat{G}(e^{j\omega}) \). From Equation 2.25, the modelling error in the process gain, obtained via the least squares solution, is given by

\[
|\Delta G(z)|_{z=e^{j\omega}} = \left| \hat{G}(z) - G(z) \right|_{z=e^{j\omega}} \leq p \left( \frac{\sigma_a^2}{\lambda_0} \right)^{\frac{1}{2}} \tag{4.22}
\]

The parameter \( p \) is a positive, real number which determines the probability of the true modelling error satisfying the bound, e.g. \( p = 1, 2, 3, \ldots \) means that the modelling error satisfies the bound with probability 0.683, 0.954, 0.997, \ldots according to the specified level of the normal distribution (Goberdansingh et al., 1992). Dividing both sides of Equation 4.22 by \( G(z) \) gives

\[
\left| \frac{\Delta G(z)}{G(z)} \right|_{z=e^{j\omega}} \leq \frac{p \sigma_a^2}{G(z) \lambda_0^{\frac{1}{2}}} \tag{4.23}
\]

The left hand side of Equation 4.23 represents the fractional error in the gain estimate \( G(e^{j\omega}) \), and will be denoted as \( \tau_0 \). Rearranging Equation 4.23 for \( \lambda_0 \), and making the relation an equality gives

\[
\lambda_0 = \lambda_0^{\text{opt}} = \left( \frac{p \sigma_a^2}{\tau_0 G(e^{j\omega})} \right)^2 \tag{4.24}
\]

With the above specification of \( \lambda_0^{\text{opt}} \), the process gain estimated from this optimal experiment will be in error of \( \tau_0 \times 100\% \) at the confidence level associated with \( p \).

Note that Equation 4.24 is a function of the true process gain \( G(e^{j\omega}) \), which is unknown. As an approximation, an estimate of the gain calculated after the first phase of the experiment will be used in place of \( G(e^{j\omega}) \). So, in addition to estimating the weighting function, noise model, and background spectrum, an estimate of the process gain will also be computed at the end of the first phase. In order to minimize the computation time, \( G(e^{j\omega}) \) will be estimated by using a FSF model with a reduced order of \( n = 1 \). Although the input-output data from the first phase of the experiment contains energy at multiple frequencies, the orthogonality of the FSF filters allows the frequency response at any given frequency \( G(e^{j\omega}) \) in this case) to be estimated independently of the other parameters in the FSF model. The calculation can also be made more efficient by manually prefiltering the input-output data with the inverse of the noise model, rather than using the computationally intensive GLS algorithm. The noise variance \( \sigma_a^2 \) will also be available from the
estimation of the noise model \( \hat{H}(z) \) as in Equation 3.26, or alternatively from the application of the GLS algorithm in the estimation of the weighting function \( Q(e^{j\omega_k}) \) as described in Section 4.2.1.

4.3.2 Design of Phase 2 Excitation

As a review up to this point, the following is a list of the quantities that are estimated at the end of the first phase:

- The amplitude response of the weighting function, \( \left| Q(e^{j\omega_k}) \right|^2 \) for \( k = 0, \ldots, n-1 \)
- The noise model, \( \hat{H}(z) \)
- The background spectrum, \( \tilde{X}_{k}^{b} \) for \( k = 0, \ldots, n-1 \)
- The process gain, \( \hat{G}(e^{j\omega_0}) \)

It will also be necessary to calculate the input signal energy at the end of the first phase in order to determine the energy requirements for the second phase. Denoted as \( \tilde{\lambda}_{k}^{1} \), this energy can be calculated by

\[
\tilde{\lambda}_{k}^{1} = \sum_{t=0}^{M_1-1} \phi^{b*}(t)\phi^{b}(t) \tag{4.25}
\]

for \( k = 0, \ldots, n-1 \), where

\[
\phi^{b}(t) = \frac{1}{N} \frac{1 - z^{-N}}{1 - e^{j\omega_k}z^{-1}} \frac{u(t)}{\hat{H}(z)} \tag{4.26}
\]

Note that the data \( u(t) \) is prefILTERED with the inverse of the estimated noise model \( \hat{H}(z) \) in Equation 4.26.

With the information acquired from the first phase of the experiment, the on-line calculation of the reference spectrum required in phase two may now be performed. The optimal reference spectrum for the second phase of the experiment is given by

\[
\lambda_{k}^{opt2} = \frac{|H(e^{j\omega_k})|^2}{|Q(e^{j\omega_k})|^2} \left( \lambda_{k}^{opt1} - \frac{M_2}{M_1} \frac{\lambda_{k}^{b1}}{|H(e^{j\omega_k})|^2} \right) \tag{4.27}
\]

for \( k = 0, \ldots, n-1 \).

To derive the result in Equation 4.27, it is sufficient to show that a two phase experiment with the reference signal in the second phase designed in accordance with Equation 4.27 is
equivalent to a single phase optimal experiment. From Equation 4.3, the total energy of the prefiltered input in the first phase of the experiment is assumed to be generated by

$$\lambda_k^1 = \frac{|Q(e^{j\omega_k})|^2 \lambda_k^{\mu p 1}}{|H(e^{j\omega_k})|^2} + \frac{\lambda_k^{/b1}}{|H(e^{j\omega_k})|^2}$$

(4.28)

where $\lambda_k^{\mu p 1}$ represents the $k$th FSF energy in the reference signal used in the first phase of the experiment. Rearranging for $\lambda_k^{\mu p 1}$ gives

$$\lambda_k^{\mu p 1} = \frac{|H(e^{j\omega_k})|^2 \lambda_k^1 - \lambda_k^{/b1}}{|Q(e^{j\omega_k})|^2}$$

(4.29)

Adding Equations 4.27 and 4.29 then gives

$$\lambda_k^{\mu p 1} + \lambda_k^{\mu p 2} = \frac{|H(e^{j\omega_k})|^2}{|Q(e^{j\omega_k})|^2} \left( \lambda_k^{\text{opt}} - \frac{\lambda_k^{/b1}}{|H(e^{j\omega_k})|^2} - \frac{\lambda_k^{/b2}}{M_2 M_1 |H(e^{j\omega_k})|^2} \right)$$

(4.30)

Noting Equation 4.20, the last term of Equation 4.30 can be simplified by

$$\frac{M_2}{M_1 |H(e^{j\omega_k})|^2} = \frac{\lambda_k^{/b2}}{|H(e^{j\omega_k})|^2}$$

(4.31)

Substituting Equation 4.31 into Equation 4.30, and rearranging gives

$$\lambda_k^{\mu p 1} + \lambda_k^{\mu p 2} = \frac{|H(e^{j\omega_k})|^2 \lambda_k^{\text{opt}} - \lambda_k^{/b1} - \lambda_k^{/b2}}{|Q(e^{j\omega_k})|^2}$$

(4.32)

Assume the following two relations hold:

$$\lambda_k^{\mu p 1} + \lambda_k^{\mu p 2} = \lambda_k^{\mu p}$$

(4.33)

$$\lambda_k^{/b1} + \lambda_k^{/b2} = \lambda_k^b$$

(4.34)

Equation 4.32 can then be simplified to give

$$\lambda_k^{\mu p} = \frac{|H(e^{j\omega_k})|^2 \lambda_k^{\text{opt}} - \lambda_k^b}{|Q(e^{j\omega_k})|^2}$$

(4.35)

which is identical to Equation 4.10, thus ending the derivation.

Upon completion of the second phase of the experiment, the data collected from both phases will be used in off-line estimations of the process and noise models. With computation time not being a major concern at this point, the GLS algorithm may be used to make simultaneous estimates of $G(z)$ and $H(z)$. This is in contrast to using the averaging routine to first estimate a noise model, which can then be used as a prefilter prior to least
squares estimation of \( G(z) \). Therefore, there are no restrictions on how the excitation signal is implemented in the second phase of the experiment. That is, the reference signal does not necessarily have to be periodic. Alternatively, the van den Bos algorithm may in fact be used to compute a single period of length \( M_2 \) (which should be an integer multiple of \( N \)) which will have the FSF spectrum given in Equation 4.27.

While computation time in off-line calculations may not be a major concern, careful consideration must still be given to how efficiently the calculations between the two phases are made. With respect to the design of the second phase reference signal, the van den Bos algorithm will clearly require more time to calculate a single period of length \( M_2 \) than it would for for some shorter period length. Therefore, implementing the reference signal as a periodic waveform is favourable. As in the first phase, it would seem that the best choice for the period of excitation \( p_2 \) is the process settling time \( N \).

### 4.4 Simulation Studies

Presented here is a series of simulation studies that will illustrate the use of the proposed two phase approach for optimal closed-loop identification. The first of three studies is used simply as an illustration of how the proposed approach is implemented on a given closed-loop system. This is followed by a comparative example in which the system is identified using a single phase experiment. The third study is a series of MonteCarlo simulations investigating the reproducibility of the two phase approach in achieving a desired FSF spectrum.

#### 4.4.1 A Step-by-Step Example

A straightforward application of the proposed two phase method is presented here. The objective of this section is to give a step-by-step description of how the two phase approach is implemented on a given closed-loop system. Where appropriate, estimated quantities are compared to their true values, as a means for tracking potential sources of error in the experiment design. In this study, the true system being identified is

\[
y(t) = \frac{0.43z^{-3}}{1 - 0.92z^{-1}}u(t) + \frac{1}{1 - 0.5z^{-1}}a(t)
\]  

(4.36)
where $\sigma_a^2 = 1$ and a Proportional-Integral-Derivative (PID) controller is used as the feedback mechanism. The control law is given by

$$K(z) = 0.3 \frac{1 - 1.5z^{-1} + 0.55z^{-2}}{1 - z^{-1}}$$  \hspace{1cm} (4.37)

The settling time of the open-loop process given in Equation 4.36 is 284 seconds. Therefore, with the sampling interval chosen to be 4 seconds, $N$ is equal to 71. Based on recommendations made in Sections 4.2.4 and 4.3.2, the period length used in each phase, $P_1$ and $P_2$, will be 71 samples. Each phase is composed of 10 full periods of data. That is, $M_1$ and $M_2$ are chosen to be 710 samples, and therefore the total experiment length is 1420 samples. It is assumed that there is no constraint on the amplitude of the test signal.

**Step 1: Design of Reference Signal for Phase One**

Refer to the procedure given in Section 2.6 for multisine signal design. In situations where there is a constraint on the amplitude of the reference signal, the desired spectrum may need to be specified in a trial-and-error fashion as described in Section 4.2.4. In this simulation, there are no constraints on the signal, so the desired reference signal spectrum is arbitrarily taken as $\lambda_k^{\text{ref}} = 750$ for $k = 0, \ldots, \frac{n-1}{2}$. It should be noted that the parameters $P$ and $M$ in the second step of the van den Bos algorithm are taken as $P_1$ and $M_1$, respectively. The computed reference signal, which has a peak amplitude of 9.651, is presented in Figure 4.2.

**Step 2: Execution of Phase One**

The reference signal computed in Step 1 above is used to excite the closed-loop system. Measurements of the process input and output are taken, and are presented in Figure 4.3.

**Step 3: Calculations at End of Phase One**

The calculations made in Step 3 correspond to those that would be made on-line between the two phases of the experiment.

**Step 3A: Estimation of the Weighting Function $Q(e^{i\omega_k})$**

In Step 3A, the frequency response of the weighting function $Q(z)$ at the frequencies $\omega_k$ for $k = 0, \ldots, \frac{n-1}{2}$ is identified. A FSF model with $N = 71$ and $n = 13$ is estimated from
measurements of \( y_{sp}(t) \) and \( u(t) \) using a GLS algorithm. The estimated frequency response coefficients and the corresponding amplitude response of \( Q(z) \) are presented in Table 4.1. It is also interesting to note that the GLS algorithm determined the noise \( \hat{H}_n(z, \rho) \) in Equation 4.12 to be white.

![Figure 4.2: Phase One Reference Signal](image)

| \( k \) | \( Q(e^{j\omega_k}) \) | \( Q'(e^{j\omega_k}) \) | \( |Q(e^{j\omega_k})|^2 \) | \( |Q'(e^{j\omega_k})|^2 \) |
|---|---|---|---|---|
| 0 | 0.1860 | 0.1767 | 0.0346 | 0.0312 |
| 1 | 0.1947 - 0.0254j | 0.1941 - 0.0322j | 0.0386 | 0.0387 |
| 2 | 0.1531 - 0.0113j | 0.1463 - 0.0053j | 0.0236 | 0.0214 |
| 3 | 0.1438 + 0.0166j | 0.1549 + 0.0297j | 0.0210 | 0.0249 |
| 4 | 0.1447 + 0.0401j | 0.1379 + 0.0393j | 0.0226 | 0.0206 |
| 5 | 0.1502 + 0.0604j | 0.1593 + 0.0609j | 0.0262 | 0.0291 |
| 6 | 0.1586 + 0.0784j | 0.1649 + 0.0793j | 0.0313 | 0.0335 |

Table 4.1: Phase One Estimate of the Weighting Function

**Step 3B: Estimation of the Noise Model \( H(z) \)**

The averaging routine is used to estimate the noise model. The noise is modelled as an autoregressive sequence, with the model order being determined by the PRESS statistic.
Figure 4.3: Phase One Input-Output Measurements

(Wang and Cluett, 1996). From measurements of the process output $y(t)$, the noise model is estimated as

$$\hat{H}(z) = \frac{1}{1 - 0.489z^{-1}}$$

(4.38)

It is also necessary to calculate the amplitude response of $\hat{H}(z)$, which has been included in Table 4.2, along with the amplitude response of the true noise model.

| $k$ | $|H(e^{j\omega_k})|^2$ | $|\hat{H}(e^{j\omega_k})|^2$ |
|-----|---------------------|---------------------|
| 0   | 4.0000              | 3.8294              |
| 1   | 3.9384              | 3.7740              |
| 2   | 3.7647              | 3.6177              |
| 3   | 3.5083              | 3.3852              |
| 4   | 3.2051              | 3.1078              |
| 5   | 2.8876              | 2.8143              |
| 6   | 2.5794              | 2.5266              |

Table 4.2: Phase One Estimate of the Noise Model

**Step 3C: Estimation of the Background Spectrum $\lambda^{bl}_k$**

As described in Section 4.2.3, the averaging routine is also used to estimate the background
spectrum, \( \lambda_k^{(b)} \) for \( k = 0, \ldots, \frac{m-1}{2} \). As a basis for comparison, the 'true' background spectrum can be determined by running the simulation without external excitation, in which case the measured input \( u(t) \) will be generated by the feedback only. The results are presented in Table 4.3.

\[
\begin{array}{|c|c|c|}
\hline
k & \lambda_k^{(b)} & \lambda_k^{(1)} \\
\hline
0 & 1.1242 & 0.9995 \\
1 & 1.5592 & 1.5082 \\
2 & 0.7708 & 0.5910 \\
3 & 1.0052 & 0.7200 \\
4 & 0.5997 & 0.4811 \\
5 & 0.6727 & 0.4963 \\
6 & 0.6085 & 0.4990 \\
\hline
\end{array}
\]

Table 4.3: Phase One Estimate of the Background Spectrum

**Step 3D: Estimation of Phase One FSF Spectrum \( \lambda_k^1 \)**

Equation 4.25 is used to calculate the energy present in the input signal at the end of the first phase. The data, which is presented in Table 4.4, seems to follow the expected trend. Specifically, since the reference spectrum in the first phase is flat and the noise \( \tilde{H}_u(z, \rho) \) is white, the distribution of energy in the input \( u(t) \) should follow a trend similar to that of \( |\tilde{Q}(e^{j\omega_k})|^2 \) according to Equation 4.28.

\[
\begin{array}{|c|c|}
\hline
k & \lambda_k^1 \\
\hline
0 & 6.7671 \\
1 & 8.3104 \\
2 & 4.5570 \\
3 & 5.6459 \\
4 & 5.2316 \\
5 & 7.9243 \\
6 & 10.3069 \\
\hline
\end{array}
\]

Table 4.4: Phase One Input Energy

**Step 3E: Calculation of the Desired FSF Spectrum \( \lambda_k^{opt} \)**

The issue of specifying a desired energy level was discussed in Section 4.3.1. For this study, the tolerable error \( \tau_0 \) in the gain estimate \( \tilde{G}(e^{j\omega}) \) is chosen as 2.5% at a confidence level of 95% (\( p = 2 \)). With the noise variance \( \sigma_n^2 \) estimated as 0.99, and the gain estimate being
5.71, the desired total energy at the zero frequency $\lambda^\text{opt}_0$ is 194.09 (see Equation 4.24). The desired energy at the other FSF frequencies may now be set in accordance with the optimal distribution given in Table 2.1. The calculated $\lambda^\text{opt}_k$ for $k = 0, \ldots, \frac{n-1}{2}$ are presented in Table 4.5.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\lambda^\text{opt}_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>194.0946</td>
</tr>
<tr>
<td>1</td>
<td>114.1733</td>
</tr>
<tr>
<td>2</td>
<td>57.0867</td>
</tr>
<tr>
<td>3</td>
<td>38.8189</td>
</tr>
<tr>
<td>4</td>
<td>29.4083</td>
</tr>
<tr>
<td>5</td>
<td>23.6701</td>
</tr>
<tr>
<td>6</td>
<td>20.0098</td>
</tr>
</tbody>
</table>

Table 4.5: Specification of the Desired FSF Spectrum

**Step 3F: Calculation of the Required Phase Two Reference Spectrum $\lambda^{\text{ref},2}_k$**

With the information gathered in Tables 4.1 to 4.5, the energy required in the reference signal over the second phase of the experiment may now be calculated via Equation 4.37.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\lambda^{\text{ref},2}_k$</th>
<th>$\lambda^{\text{true}}_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>22949</td>
<td>21626</td>
</tr>
<tr>
<td>1</td>
<td>10282</td>
<td>10769</td>
</tr>
<tr>
<td>2</td>
<td>8832</td>
<td>8360</td>
</tr>
<tr>
<td>3</td>
<td>4484</td>
<td>5506</td>
</tr>
<tr>
<td>4</td>
<td>3632</td>
<td>3409</td>
</tr>
<tr>
<td>5</td>
<td>1507</td>
<td>1710</td>
</tr>
<tr>
<td>6</td>
<td>717</td>
<td>780</td>
</tr>
</tbody>
</table>

Table 4.6: Desired Phase Two Reference Spectrum

In Table 4.6 it is clear that a larger portion of the energy in the reference signal will be concentrated in the low frequency region. As a basis for comparison, $\lambda^{\text{true}}_k$ was also calculated from Equation 4.27 using the true values of $|Q(e^{j\omega_k})|^2$, $|H(e^{j\omega_k})|^2$, and $\lambda^1_k$.

**Step 3G: Design of Reference Signal for Phase Two**

Step 1 is repeated to design a reference signal consisting of the FSF energies $\lambda^{\text{ref},2}_k$ listed in Table 4.6. As in the first phase, the reference signal in the second phase will be composed of 10 periods of a multisine with a period length equal to 71. The computed reference signal,
which has a peak amplitude of 19.94, is presented in Figure 4.4. The design of this reference signal marks the end of the first phase. The time required to perform all calculations in Step 3 was less than 1.54 seconds\(^2\), which is considerably less than the sampling interval of 4 seconds.

**Step 4: Execution of Phase Two**

The second phase of the experiment is started at time \( t = M_1 \) by injecting the calculated reference signal into the closed-loop system. The experiment ends at time \( t = M - 1 \). The measurements taken throughout the full experiment are presented in Figure 4.5.

**Step 5: Completion of the Two-Phase Experiment**

With the experiment completed, off-line calculations can now be made. A GLS algorithm was used to estimate a process model \( \hat{G}(z) \) and a noise model \( \hat{H}(z) \). Of particular interest is how closely the actual FSF spectrum \( \lambda_k \) matches the desired spectrum \( \lambda_k^{\text{opt}} \) specified in Table 4.5. The desired and actual FSF spectra are plotted in Figure 4.6. The results are

\(^2\)On an Intel Celeron 400 processor, 128 MB RAM.
also summarized in Table 4.7. It seems that the actual FSF spectrum approximates the desired spectrum reasonably well over the entire frequency range. In addition to the spectrum, a FSF-generated estimate of the process step response with 95% confidence bounds was computed using Equations 2.22 and 2.28, respectively. The step response estimate is presented in Figure 4.7. The process gain was estimated to be \(5.49 \pm 2.7\%\) at a confidence level of 95%, which has a somewhat greater relative error \(\tau_0\) than the specified 2.5%. The noise model estimated by the GLS algorithm is

\[
\hat{H}(z) = \frac{1}{1 - 0.4968z^{-1}}
\]

which is an improvement over the estimate made after the first phase of the experiment. Also noteworthy in Table 4.7 is that perfect knowledge of \(|Q(e^{j\omega_k})|^2\), \(|H(e^{j\omega_k})|^2\), and \(\lambda_k^{\text{opt}}\) somewhat improved the results, as evidenced by the values of \(\lambda_k^{\text{true}}\) and \(\psi_k^{\text{true}}\). The spectrum \(\lambda_k^{\text{true}}\) is that obtained by implementing a reference signal with a spectrum of \(\lambda_k^{\text{true}}\) in the second phase of the experiment, as given in Table 4.6. Possible causes for the small difference between \(\lambda_k^{\text{opt}}\) and \(\lambda_k^{\text{true}}\) obtained from the two-phase approach will be discussed in Section 4.5.
Figure 4.6: Input Signal FSF Energy Distribution, desired (○), actual (○)

Figure 4.7: FSF-Generated Step Response Estimate: Two Phase Approach, estimate (-), true response (-), 95% confidence bounds (- -)
4.4.2 A Comparative Example

In this example, the system studied in Section 4.4.1 is identified using a single phase experiment with the same amount of total energy in the reference signal. That is,

\[
\sum_{k=0}^{6} \lambda_k^{\text{opt}} = \sum_{k=0}^{6} \left( \lambda_k^{\text{opt}^1} + \lambda_k^{\text{opt}^2} \right)
\]  

However, in this example the energy in the reference signal is equally distributed among the FSF frequencies. The length of the experiment is 1420 samples. with \( \lambda_k^{\text{opt}} = 8237.14 \) for \( k = 0, \ldots, 141 \). A periodic signal is used, and as in the previous section, the period length is set to 71 samples.

The simulated input-output data is presented in Figure 4.8. The FSF/GLS algorithm was applied to estimate the noise model and process step response. The noise model was determined to be

\[
\tilde{H}(z) = \frac{1}{1 - 0.4889z^{-1}}
\]  

The step response estimate with 95% confidence bounds is shown in Figure 4.9.

In order to assess the quality of the experiment, recall the SE measure defined in Equation 2.34:

\[
SE = \sum_{m=0}^{N-1} 2\delta(m)
\]

where \( \delta(m) \) is the standard deviation of the step response estimate at the sampling instant \( m \). The SE describes the area enclosed by the 95% confidence bounds, and is minimized when the FSF spectrum energies \( \lambda_k' \) are distributed in accordance with Table 2.1. For this example, the SE value was calculated as 10.8. The SE value using the two phase approach in the previous section is significantly lower at 8.7. However, these SE values are not

<table>
<thead>
<tr>
<th>( k )</th>
<th>( \lambda_k^{\text{opt}} )</th>
<th>( \psi_k^{\text{optimal}} )</th>
<th>( \lambda_k' )</th>
<th>( \psi_k )</th>
<th>( \lambda_k^{\text{true}} )</th>
<th>( \psi_k^{\text{true}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>194.0946</td>
<td>1.0</td>
<td>191.6119</td>
<td>1.0</td>
<td>183.2302</td>
<td>1.0</td>
</tr>
<tr>
<td>1</td>
<td>114.1733</td>
<td>1.7</td>
<td>105.7805</td>
<td>1.8</td>
<td>108.7442</td>
<td>1.7</td>
</tr>
<tr>
<td>2</td>
<td>57.0867</td>
<td>3.4</td>
<td>59.8921</td>
<td>3.2</td>
<td>56.9048</td>
<td>3.2</td>
</tr>
<tr>
<td>3</td>
<td>38.8189</td>
<td>5.0</td>
<td>33.6814</td>
<td>5.7</td>
<td>39.6088</td>
<td>4.6</td>
</tr>
<tr>
<td>4</td>
<td>29.4083</td>
<td>6.6</td>
<td>30.3244</td>
<td>6.3</td>
<td>29.7560</td>
<td>6.2</td>
</tr>
<tr>
<td>5</td>
<td>23.6701</td>
<td>8.2</td>
<td>19.8092</td>
<td>9.7</td>
<td>24.1707</td>
<td>7.6</td>
</tr>
<tr>
<td>6</td>
<td>20.0098</td>
<td>9.7</td>
<td>20.3486</td>
<td>9.4</td>
<td>19.4674</td>
<td>9.4</td>
</tr>
</tbody>
</table>

Table 4.7: Comparison of the Desired and Actual FSF Spectra

4.4.2 A Comparative Example
Figure 4.8: Input-Output Measurements

Figure 4.9: FSF-Generated Step Response Estimate, estimate (-), true response (-), 95% confidence bounds (- -)
directly comparable since the total energy in the input differs for the two cases. Table 4.8 presents the FSF spectrum $\lambda_k$ for this example along with the optimal FSF spectrum $\lambda_k^{opt}$ that would minimize the $SE$ value with an equal amount of total input energy. That is, in Table 4.8,

$$\sum_{k=0}^{6} \lambda_k = \sum_{k=0}^{6} \lambda_k^{opt} \tag{4.43}$$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\lambda_k$</th>
<th>$\lambda_k^{opt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>71.3322</td>
<td>209.0591</td>
</tr>
<tr>
<td>1</td>
<td>87.6356</td>
<td>122.9760</td>
</tr>
<tr>
<td>2</td>
<td>57.6409</td>
<td>61.4880</td>
</tr>
<tr>
<td>3</td>
<td>52.1341</td>
<td>41.8118</td>
</tr>
<tr>
<td>4</td>
<td>61.4088</td>
<td>31.6736</td>
</tr>
<tr>
<td>5</td>
<td>79.1868</td>
<td>25.4950</td>
</tr>
<tr>
<td>6</td>
<td>104.3203</td>
<td>21.5525</td>
</tr>
</tbody>
</table>

Table 4.8: FSF Spectrum Energies

From Table 4.8, it is clear that the experiment was suboptimal. As a measure of the optimality of the experiment, the $SE$ values may be compared to the theoretical minimum for the simulations presented in this and the previous section, as follows:

$$SE_{min} = \sum_{m=0}^{N-1} 2\delta_{min}(m) \tag{4.44}$$

where

$$\delta_{min}(m)^2 = W(m)(\Phi^\ast \Phi)^{-1}_{opt}W(m)\sigma_a^2 \tag{4.45}$$

The vector $W(m)$ appears in Equation 2.26. The matrix $(\Phi^\ast \Phi)_{opt}$ represents the correlation matrix defined in Equation 2.29, with $\nu_{kk,kk}^{opt} = \lambda_k^{opt}$ for $k = 0, \ldots, n-1$. The off-diagonal elements of $(\Phi^\ast \Phi)_{opt}$ are set to zero in Equation 4.45. Table 4.9 compares the $SE$ and $SE_{min}$ values. The $SE_{min}$ values are based on the optimal spectra $\lambda_k^{opt}$ in Tables 4.5 and 4.8 for the two phase and single phase approaches, respectively.

<table>
<thead>
<tr>
<th>Case</th>
<th>$SE_{min}$</th>
<th>$SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two Phase Approach</td>
<td>8.6</td>
<td>8.7</td>
</tr>
<tr>
<td>Single Phase Approach</td>
<td>8.2</td>
<td>10.8</td>
</tr>
</tbody>
</table>

Table 4.9: Optimal and Achieved $SE$ Values
Application of the two phase approach significantly outperformed the single phase approach. The SE value of 8.7 calculated for the two approach is much closer to its corresponding $SE_{\min}$ value of 8.6, in contrast to the single phase approach (which had an SE value of 10.8 compared to a $SE_{\min}$ value of 8.2). With a flat reference spectrum being used in this example, a disproportionate amount of energy was passed into the higher frequencies of the input signal, most notably because of the shape of $|Q(e^{j\omega k})|^2$ for $k = 0, \ldots, \frac{n-1}{2}$ given in Table 4.1. Consequently, the input spectrum $\lambda_k$ does not even remotely resemble the optimal shape, resulting in a high SE value.

4.4.3 Monte Carlo Simulations

In this study, the true system being identified is a 2nd order, overdamped process given by

$$y(t) = \frac{0.044z^{-1} + 0.038z^{-2}}{1 - 1.62z^{-1} + 0.64z^{-2}}u(t) + \frac{1}{1 - 0.90z^{-1}}a(t)$$  \hspace{1cm} (4.46)

where $\sigma_a^2 = 1$. The Proportional-Integral (PI) control law is given by

$$K(z) = \frac{0.01}{1 - z^{-1}}$$  \hspace{1cm} (4.47)

The settling time of the open-loop process given in Equation 4.46 is 360 seconds. Therefore, with the sampling interval chosen to be 4 seconds, $N$ is equal to 90. Based on recommendations made in Sections 4.2.4 and 4.3.2, the period length used in each phase, $P_1$ and $P_2$, will be 90 samples. Each phase is composed of 10 full periods of data. That is, $M_1$ and $M_2$ are chosen to be 900 samples, and therefore the total experiment length is 1800 samples. It is assumed that there is no constraint on the amplitude of the test signal.

The reference signal used in the first phase of the experiment is a multisine with its energy distributed as in Table 2.1, where $\lambda_0^{ref} = 3000$. This reference signal was applied to the closed-loop system in a series of simulations consisting of 250 different realizations of the noise $a(t)$. Two different options are explored at the end of the first phase. One set of results is generated from the use of the GLS in the estimation of the weighting function $\hat{Q}(e^{j\omega k})$ for $k = 0, \ldots, \frac{n-1}{2}$, while the other set applies a standard least squares (SLS) estimator without modelling $\hat{H}_u(z, \rho)$. Rather than selecting a tolerable error $\tau_0$ in the gain estimate, a single desired FSF spectrum $\lambda_k^{opt}$ for $k = 0, \ldots, \frac{n-1}{2}$ was used for all simulations.

Table 4.10 summarizes the results of estimating the weighting function $|Q(e^{j\omega k})|^2$ by use of the GLS algorithm. The table compares the true value of $|Q(e^{j\omega k})|^2$ to the average
of 250 estimates, as well as the minimum and the maximum value of 250 estimates. Of particular interest is the wide range of values estimated at the zero frequency. There is also a significant error in the average value \(|\bar{Q}(e^{j\omega_0})|^2_{\text{avg}}\).

\[
\begin{array}{|c|c|c|c|c|}
\hline
k & |Q(e^{j\omega_k})|^2 \times 10^4 & \bar{Q}(e^{j\omega_k})|^2_{\text{avg}} \times 10^4 & \bar{Q}(e^{j\omega_k})|^2_{\text{min}} \times 10^4 & \bar{Q}(e^{j\omega_k})|^2_{\text{max}} \times 10^4 \\
\hline
0 & 594.9 & 1133.0 & 0.0005621 & 7285 \\
1 & 417.6 & 430.5 & 152.6 & 886.5 \\
2 & 65.77 & 68.27 & 29.56 & 128.2 \\
3 & 25.35 & 25.82 & 11.48 & 50.73 \\
4 & 13.57 & 13.83 & 6.042 & 24.35 \\
6 & 5.875 & 5.904 & 3.058 & 9.708 \\
\hline
\end{array}
\]

Table 4.10: Phase One GLS Estimate of the Weighting Function

The high variance in the estimates of \(|Q(e^{j\omega_0})|^2\) can be attributed to the low amount of energy at the zero frequency after prefiltering the data with the inverse of the noise model \(\tilde{H}_u(z, \rho)\). The estimated noise model \(\tilde{H}_u(z, \rho)\) contains a pole very close to the unit circle, and therefore the consequent prefiltering operation removes over 99% of the energy from the zero frequency. That is, the energy used in estimating \(|Q(e^{j\omega_0})|^2\) is reduced from 3000 to approximately 5.0 as a result of the prefiltering.

The accuracy of the identified weighting function has a significant effect on the overall performance of the two-phase approach in obtaining a desired FSF spectrum \(\lambda_k^{\text{opt}}\) for \(k = 0, \ldots , \frac{n-1}{2}\). Table 4.11 presents the average, minimum, and maximum FSF energy obtained at each frequency from 250 simulations. Also presented in Table 4.11 is \(\lambda_k^{\text{true}}\), the average FSF spectrum obtained when the true values of \(|Q(e^{j\omega_0})|^2\), \(|H(e^{j\omega_0})|^2\), and \(\lambda_k^{\text{blue}}\) are used in Equation 4.27 to calculate \(\lambda_k^{\text{true}}\) for \(k = 0, \ldots , \frac{n-1}{2}\).

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
k & \lambda_k^{\text{opt}} & \lambda_k^{\text{true}} & \lambda_k^{\text{avg}} & \lambda_k^{\text{min}} & \lambda_k^{\text{max}} \\
\hline
0 & 78.4720 & 87.0746 & 4485.48 & 10.2968 & 111544801 \\
1 & 46.1600 & 48.6684 & 2845.03 & 24.2056 & 693152 \\
2 & 23.0800 & 23.0842 & 462.529 & 13.0184 & 108711 \\
3 & 15.6944 & 15.2814 & 190.097 & 8.0276 & 43191 \\
4 & 11.8897 & 11.4350 & 106.218 & 6.5816 & 23441 \\
5 & 9.5698 & 9.1317 & 69.1802 & 5.4921 & 14878 \\
6 & 8.0899 & 7.6283 & 49.3944 & 4.6425 & 10301 \\
\hline
\end{array}
\]

Table 4.11: FSF Spectrum Energies Based on GLS Estimates of the Weighting Function
The wide range of energies observed at each FSF frequency is a direct result of the high variance in the estimated weighted function. It would seem that the application of the GLS to the identification of the weighting function is in fact detrimental to the performance of the two-phase approach. As a comparison, Table 4.12 presents the results of applying a SLS algorithm to the estimation of $|Q(e^{j\omega_k})|^2$. The average estimates are reasonably accurate and significantly improved over those presented in Table 4.10. Furthermore, the range of values estimated for $|Q(e^{j\omega_k})|^2$ is reduced, particularly at the lower FSF frequencies.

| $k$ | $|Q(e^{j\omega_k})|^2 \times 10^4$ | $\bar{Q}(e^{j\omega_k})^2_{\text{avg}} \times 10^4$ | $\bar{Q}(e^{j\omega_k})^2_{\text{min}} \times 10^4$ | $\bar{Q}(e^{j\omega_k})^2_{\text{max}} \times 10^4$ |
|-----|-------------------------------|-----------------------------|-----------------|-----------------------------|
| 0   | 594.9                         | 592.5                       | 87.75           | 1252.2                     |
| 1   | 417.6                         | 432.2                       | 135.2           | 901.3                      |
| 2   | 65.77                         | 68.99                       | 28.53           | 129.6                      |
| 3   | 25.35                         | 25.70                       | 11.35           | 54.28                      |
| 4   | 13.57                         | 14.10                       | 5.566           | 25.47                      |
| 5   | 8.519                         | 8.896                       | 3.989           | 15.56                      |
| 6   | 5.875                         | 5.936                       | 2.953           | 9.450                      |

Table 4.12: Phase One SLS Estimate of the Weighting Function

Table 4.13 summarizes the results obtained from the two-phase approach based on SLS estimates of the weighting function. The family of 250 FSF spectra is presented in Figure 4.10, while the average spectrum is compared to the desired spectrum $\lambda_k^\text{det}$ for $k = 0, \ldots, n-1$ in Figure 4.11. From Figure 4.10 it is clear that the energy at the FSF frequencies corresponding to $k = 0$ and $k = 1$ tend to be greater than the desired values for this closed-loop system. Nevertheless, the average FSF spectrum based on SLS estimates of the weighting function shows a marked improvement over that obtained from the GLS method. Furthermore, the average FSF spectrum closely approximates the spectrum that can be obtained by using the true values of $|Q(e^{j\omega_k})|^2$, $|H(e^{j\omega_k})|^2$, and $\lambda_k^{01}$ in Equation 4.27.

The results of this section strongly suggest that modelling the noise $\tilde{H}_x(z, \rho)$ adversely affects the optimality of the two-phase approach when applied to the specific closed-loop system being studied. In conditions where the filter $H(z)S(z)K(z)$ in Equation 4.11 does not contain poles near the unit circle, the impact of applying a GLS algorithm to the estimation of the weighting function may not be as significant. As discussed in Section 4.2.1, since the weighting function $Q(e^{j\omega_k})$ can be estimated consistently without modelling the noise, it would seem reasonable to completely eliminate the use of a GLS algorithm.
Figure 4.10: Input Signal FSF Energy Distribution

Figure 4.11: Average Input Signal FSF Energy Distribution, desired (o), average (o)
Table 4.13: FSF Spectrum Energies Based on SLS Estimates of the Weighting Function

<table>
<thead>
<tr>
<th>k</th>
<th>$\lambda_k^{opt}$</th>
<th>$\lambda_k^mtrwc$</th>
<th>$\lambda_k^{avg}$</th>
<th>$\lambda_k^{min}$</th>
<th>$\lambda_k^{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>78.4720</td>
<td>86.7549</td>
<td>125.9128</td>
<td>34.5622</td>
<td>519.2892</td>
</tr>
<tr>
<td>1</td>
<td>46.1600</td>
<td>48.3715</td>
<td>58.2579</td>
<td>22.4958</td>
<td>154.2469</td>
</tr>
<tr>
<td>2</td>
<td>23.0800</td>
<td>23.0427</td>
<td>25.2721</td>
<td>12.6812</td>
<td>50.8475</td>
</tr>
<tr>
<td>3</td>
<td>15.5944</td>
<td>15.2952</td>
<td>16.7309</td>
<td>7.6684</td>
<td>34.9604</td>
</tr>
<tr>
<td>6</td>
<td>8.0899</td>
<td>7.6210</td>
<td>7.9735</td>
<td>4.7507</td>
<td>15.5150</td>
</tr>
</tbody>
</table>

in this stage of the proposed two-phase approach. The set of Monte Carlo simulations in which a SLS algorithm was used in the estimation of $Q(e^{j\omega_k})$ have also confirmed the reproducibility of the two-phase approach in achieving a desired FSF spectrum. While there may be some variability in FSF energies, the spectrum obtained via the two-phase approach does approach the desired spectrum within a reasonable degree of accuracy.

4.5 Possibilities and Limitations

When applying the proposed two-phase approach to optimal closed-loop identification, it is important to know the quality of results that can be expected. This section addresses the effects of the discontinuity in the reference signal between phase one and two and the use of finite-length data sets on the optimality of the two-phase experiment. In so doing, some of the assumptions made in the development of the proposed two-phase approach will be examined. The section also presents some modifications to the two-phase approach that are necessary if the noise is non-stationary.

4.5.1 The Transition Point

The transition point refers to that part of the two-phase experiment at which the first phase ends and the second phase begins. The transition point divides the reference signal into two parts, each of which is periodic. However, the reference signal $y_{sp}(t)$ when considered as a whole for $t = 0, \ldots, M - 1$ is not periodic. The discontinuity in the reference signal causes a small quantity of the total energy to be distributed or leak into frequencies other than the FSF frequencies.

The leakage of energy into non-FSF frequencies will first be illustrated by analyzing
Figure 4.12: Fourier Amplitude Spectrum of the First Phase Reference Signal

The reference signal used for the simulation study in Section 4.4.1. The Fourier amplitude spectrum $|Y_{sp}(e^{j\omega r})|$ for the first and second phase reference signals were computed, and are presented in Figures 4.12 and 4.13, respectively. In addition, the Fourier amplitude spectrum of the whole reference signal is presented in Figure 4.14.

The Fourier amplitude spectrum presented in Figure 4.12 illustrates the energy content of the reference signal used in the first phase of the experiment. The spectrum is non-zero for the harmonic numbers $r$ that correspond to the FSF frequencies $\omega_k$ for $k = 0, \ldots, \frac{n-1}{2}$, and zero elsewhere. The shape of the Fourier amplitude spectrum also reflects the use of an equal amount of energy at each of the FSF frequencies in the first phase of the experiment. Figure 4.13 is similar in that it is non-zero at the FSF frequencies only, but in this case the spectrum is not uniform since the energy is distributed in accordance with $\lambda_k^{\text{sp}}$ in Table 4.6.

The reference signal used in each phase of the experiment is periodic, with energy only

---

3See Appendix A for an explanation of the relation between the DFT harmonic number $r$ and the FSF frequencies $k$. 

---
Figure 4.13: Fourier Amplitude Spectrum of the Second Phase Reference Signal

Figure 4.14: Fourier Amplitude Spectrum of the Entire Reference Signal
in the FSF frequencies. However, the Fourier amplitude spectrum of the reference signal taken as whole, as presented in Figure 4.14, clearly indicates that there is a portion of energy in the non-FSF frequencies as well. The observed leakage of energy into the non-FSF frequencies occurs due to the discontinuity in the reference signal at the transition point. The transition point thus affects the two-phase experiment by reducing the FSF energies \( \lambda'_k \) relative to the desired values \( \lambda'^{opt}_k \). In addition, the observed leakage of energy into non-FSF frequencies would suggest that the assumption made in Equation 4.33 is not necessarily correct, introducing an error in Equation 4.27.

To understand the effect of the transition point further, consider the definition of the inverse discrete Fourier transform:

\[
y_{sp}(t) = \sum_{r=0}^{M-1} Y_{sp}(r)e^{j2\pi rt/M}
\]

for \( t = 0, \ldots, M - 1 \). Equation 4.48 can be used to construct a periodic signal \( y_{sp}(t) \) of length \( M \), given a set of Fourier coefficients \( Y_{sp}(r) \) for \( r = 0, \ldots, M - 1 \). In fact, the van den Bos algorithm for multisine design uses Equation 4.48 to swap between the time and frequency domains. Since the mapping from \( Y_{sp}(r) \) to \( y_{sp}(t) \) is 1-to-1, it is clear that any signal \( y_{sp}(t) \) that satisfies Equation 4.48 given \( Y_{sp}(r) \) for \( r = 0, \ldots, M - 1 \) will be periodic.

In the proposed two-phase experiment, the Fourier coefficients are set implicitly via \( \lambda'^{opt}_k \) (see Section 2.6), but in reality Equation 4.48 cannot be satisfied since the reference signal is not periodic. Therefore, it is impossible for the achieved FSF energies \( \lambda'_k \) to be equal to the desired spectrum \( \lambda'^{opt}_k \) for \( k = 0, \ldots, \frac{n-1}{2} \).

### 4.5.2 Finite Length Data Sets

Consider the situation depicted in Figure 4.15 in which two uncorrelated signals, \( u_1(t) \) and \( u_2(t) \), are added to form \( u(t) \). That is,

\[
u(t) = u_1(t) + u_2(t)
\]

Equation 4.49 can be used as a general representation of several signals in a given closed-loop system, such as the process input given in Equation 4.1. This section addresses the following issue regarding systems that can be described by Equation 4.49:

Is the FSF spectrum of the combined signal \( u(t) \) equal to the sum of the FSF spectra of the individual signals \( u_1(t) \) and \( u_2(t) \)?
Denote by $\phi^k(t)$ the result of passing $u_1(t)$ through the $k$th FSF filter defined by Equation 2.9 for $k = 0, \ldots, \frac{n-1}{2}$, and similarly for $\phi^k_2(t)$. In addition, the filtered values of $u(t)$ are denoted by $\phi^k(t)$. The FSF energy in each of the signals $u_1(t)$, $u_2(t)$, and $u(t)$ is respectively defined by

\begin{align*}
\lambda_{k,1} &= \sum_{t=0}^{M-1} \phi^k_1(t) \phi^*_1(t) \\
\lambda_{k,2} &= \sum_{t=0}^{M-1} \phi^k_2(t) \phi^*_2(t) \\
\lambda_k &= \sum_{t=0}^{M-1} \phi^k(t) \phi^*_k(t)
\end{align*}

Stated mathematically, the purpose of this section is to verify whether the following is true or not:

\begin{equation}
\lambda_k = \lambda_{k,1} + \lambda_{k,2}
\end{equation}

for $k = 0, \ldots, \frac{n-1}{2}$. First, substitute Equation 4.49 into Equation 2.8, giving

\begin{align*}
\phi^k(t) &= F^k(z)u_1(t) + F^k(z)u_2(t) \\
&= \phi^k_1(t) + \phi^k_2(t)
\end{align*}

Equation 4.55 may then be substituted into Equation 4.52:

\begin{equation}
\lambda_k = \sum_{t=0}^{M-1} [\phi^k_1(t) + \phi^k_2(t)]^* [\phi^k_1(t) + \phi^k_2(t)]
\end{equation}
\[ M^{-1} = \sum_{t=0}^{M-1} [\phi_1^k(t)\phi_1^k(t) + \phi_2^k(t)\phi_2^k(t) + \phi_1^k(t)\phi_2^k(t) + \phi_2^k(t)\phi_1^k(t)] \quad (4.57) \]

\[ = \lambda_{k,1} + \lambda_{k,2} + \sum_{t=0}^{M-1} [\phi_1^k(t)\phi_2^k(t) + \phi_2^k(t)\phi_1^k(t)] \quad (4.58) \]

The summation in Equation 4.58, which will be referred to as the 'cross-spectrum', is generally non-zero and therefore Equation 4.53 is false. That is, in general,

\[ \lambda_k \neq \lambda_{k,1} + \lambda_{k,2} \quad (4.59) \]

So, the FSF spectrum of the combined signal \( u(t) \) is not equal to the sum of the FSF spectra of the individual signals \( u_1(t) \) and \( u_2(t) \), and this is why Equation 4.3 was written as an approximation, and not an equality. However, it is important to know how good an approximation Equation 4.3 actually is, and if there are conditions under which it does become an equality.

**Lemma 1** For any two uncorrelated signals \( u_1(t) \) and \( u_2(t) \), the upper bound on the cross-spectrum can be described as follows:

\[ \left| \sum_{t=0}^{M-1} [\phi_1^k(t)\phi_2^k(t) + \phi_2^k(t)\phi_1^k(t)] \right| \leq 2\sqrt{\lambda_{k,1}\lambda_{k,2}} \quad (4.60) \]

See Appendix C for a proof.

The error made in assuming that the spectra in Equation 4.3 are additive may now be quantified:

\[ \left| \lambda_k - \left( |Q(e^{j\omega_k})|^2 \frac{\lambda_{kp}}{\lambda_k^2} + \frac{\lambda_{fb}}{\lambda_k^2} \right) \right| \leq 2\sqrt{|Q(e^{j\omega_k})|^2 \lambda_{kp} \lambda_{fb}} \quad (4.61) \]

More importantly, Equation 4.60 may be used for a posteriori calculations of an acceptable level of sub-optimality in an FSF identification experiment. Rearrange Equation 4.10 as follows:

\[ \lambda_k^{\text{opt}} = \frac{|Q(e^{j\omega_k})|^2 \lambda_{kp}}{|H(e^{j\omega_k})|^2 \lambda_k} + \frac{\lambda_{fb}}{|H(e^{j\omega_k})|^2} \quad (4.62) \]

Equation 4.62 represents the FSF spectrum of a summation of two signals. The following then holds:

\[ \left| \lambda_k^{\text{opt}} - \lambda_k^f \right| \leq 2\sqrt{\frac{|Q(e^{j\omega_k})|^2}{|H(e^{j\omega_k})|^2} \lambda_k \lambda_{fb}} \lambda_k \quad (4.63) \]

For the two-phase approach, Equations 4.33 and 4.19 may be substituted into Equation 4.63 to give

\[ \left| \lambda_k^{\text{opt}} - \lambda_k^f \right| \leq 2\sqrt{\frac{M |Q(e^{j\omega_k})|^2}{M_1 |H(e^{j\omega_k})|^2} \left( \lambda_{kp} + \lambda_{kb} \right) \lambda_{fb} \lambda_k} \quad (4.64) \]
All of the parameters on the right hand side of Equation 4.64 are either specified or calculated during the two-phase experiment. Equation 4.64 serves as a diagnostic tool after the two-phase experiment has been completed, defining the maximum difference that should be observed between the desired and actual FSF spectra. It should be noted that Equation 4.64 does not account for the leakage of energy that occurs due to the transition point, as discussed in Section 4.5.1.

The system defined by Equation 4.1 is a special case of Equation 4.49, in which \( u_1(t) \) is a deterministic signal (a multisine), and \( u_2(t) \) is a stochastic signal (filtered white noise). An interesting property of this type of system is that the limiting cross-spectrum is zero at all frequencies \( \omega_k \) for \( k = 0, \ldots, \frac{n-1}{2} \).

**Lemma 2** Consider the summation of two uncorrelated signals \( u_1(t) \) and \( u_2(t) \) to form \( u(t) \), as in Figure 4.15. Assume that \( u_1(t) \) is a bounded deterministic signal, and that \( u_2(t) \) is zero-mean white noise passed through a stable linear filter. The following then holds:

\[
E[ \sum_{t=0}^{M-1} \phi_1^{\ast}(t) \phi_2^k(t) ] = 0 \tag{4.65}
\]

and

\[
E[ \sum_{t=0}^{M-1} \phi_2^{\ast}(t) \phi_1^k(t) ] = 0 \tag{4.66}
\]

for \( k = 0, \ldots, \frac{n-1}{2} \).

**Proof**

Equation 4.65 may also be written as

\[
E[ \sum_{t=0}^{M-1} \phi_1^{\ast}(t) \phi_2^k(t) ] = \sum_{t=0}^{M-1} E[\phi_1^{\ast}(t) \phi_2^k(t)] \tag{4.67}
\]

\[
= \sum_{t=0}^{M-1} \phi_1^{\ast}(t) E[\phi_2^k(t)] \tag{4.68}
\]

Since \( u_2(t) \) is zero-mean white noise passed through a stable linear filter, its expected value \( E[u_2(t)] \) is equal to zero. Furthermore, since the FSF filters defined in Equation 2.9 are stable and linear, the expected value \( E[\phi_2^k(t)] \) is also zero. That is,

\[
E[\phi_2^k(t)] = \frac{1}{N} \left( 1 + e^{j\omega_k z^{-1}} + e^{2j\omega_k z^{-2}} + \cdots + e^{(N-1)j\omega_k z^{-(N-1)}} \right) E[u_2(t)] \tag{4.69}
\]

\[
= 0 \tag{4.70}
\]
Therefore, Equation 4.68 reduces to
\[
\sum_{t=0}^{M-1} \phi_1^k(t) E[\phi_2^*(t)] = 0
\]  
Equation 4.71

The proof of Equation 4.66 is similar, or it may otherwise be noted that
\[
|\phi_1^*(t)\phi_2^*(t)| = |\phi_2^*(t)\phi_1(t)|
\]  
Equation 4.72

This ends the proof.

The result of Lemma 2 becomes relevant as the length of the data set \( M \) approaches infinity. With a finite number of data points the cross-spectrum will usually be non-zero, and in fact can be quite significant when compared to the sum of the spectra of the individual signals \( u_1(t) \) and \( u_2(t) \). The magnitude of the cross-spectrum is however bounded, as shown in Lemma 1.

The discussion presented in this section clarifies that for finite data sets, Equation 4.3 may indeed only be written as an approximation. However, it is important to point out that it becomes an equality for data sets of infinite length. In addition, it should be noted that the cross-spectrum was not included in Equation 4.3 because its exact value depends on the realizations of two different signals, one of which is stochastic. That is, the cross-spectrum cannot be calculated without perfect knowledge of \( a(t) \) for \( t = 0, \ldots, M - 1 \).

4.5.3 Non-Stationary Noise

Consider the situation in which the noise model has at least one pole on the unit circle. For example, the noise model given by
\[
H(z) = \frac{1}{1 - z^{-1}}
\]  
Equation 4.73

represents a noise sequence that is non-stationary in the mean. In Section 3.3, the applicability of the averaging routine to non-stationary noise sequences was commented on. It was suggested that with the appropriate pretreatment of data, the averaging routine could be applied to non-stationary data without violating the restrictions given by McKelvey (1996). However, non-stationary noise sequences present complications to the proposed two-phase approach to optimal closed-loop experiments. To understand the source of the complication with non-stationary noise, first recall Equation 4.27:
\[
\chi_k^{\text{var}} = \frac{|H(e^{j\omega_k})|^2}{|Q(e^{j\omega_k})|^2} \left( \lambda_k^{\text{opt}} - \lambda_k - M_2 \frac{\lambda_k^{b1}}{M_1 |H(e^{j\omega_k})|^2} \right)
\]  
Equation 4.74
The energy at each frequency of the reference signal is proportional to $|H(e^{j\omega})|^2$. For a noise model with a pole on the unit circle, such as the one given in Equation 4.73, the following holds:

$$
\lim_{\omega \to 0} |H(e^{j\omega})|^2 = \infty
$$

(4.75)

With $|Q(e^{j\omega})|^2$ always being finite, the consequence of Equation 4.75 is that an infinite amount of energy is required at the zero frequency, which clearly is not feasible in a practical experiment design.

The proposed two-phase approach can be modified to allow for non-stationary noise. First, it is important to understand the properties of the FSF filters and how they relate to the FSF energies. Wang and Cluett (1997) give the following expression for the FSF energies:

$$
\lambda_k = \sum_{l=0}^{M-1} \frac{|F^k(e^{j\omega_l})|^2 |U(e^{j\omega_l})|^2}{W^k(\omega_l)}
$$

(4.76)

where $\omega_l = \frac{2\pi l}{M}$ and $F^k(e^{j\omega_l})$ is the FSF filter defined by Equation 2.9. Some properties of the weighting function $|F^k(e^{j\omega_l})|^2$, also denoted as $W^k(\omega_l)$, are given in Appendix A. The weighting function $W^0$ corresponding to the zeroth FSF frequency is plotted in Figure 4.16.

Referring to Figure 4.16, the bandwidth of $W^0$ includes frequencies from $-\frac{\pi}{N}$ to $\frac{\pi}{N}$ radians/sample. Input energy in this range of frequencies will therefore contribute to the zeroth FSF energy $\lambda_0$. The drawback with the proposed two-phase approach with respect to non-stationary noise is that energy at the center frequency of the weighting function $W^0$ is used to excite the zero frequency, leading to the problems discussed above. An alternative approach would be to use excitation at some other frequency within the bandwidth of $W^0$, such as $\pm \frac{\pi}{N}$. It can be seen in Figure 4.16 that excitation at $\pm \frac{\pi}{N}$ only contributes 40.5% of the energy that it would to $\lambda_0$ if the energy was actually at the center frequency of $W^0$. The two-phase approach may then be modified by adjusting Equation 2.37 to account for the off-centered energy at $k = 0$:

$$
U'(\frac{P}{2N}) = \sqrt{\frac{\lambda_0}{2 \times 0.405M}}
$$

(4.77)

It will therefore be the DFT harmonic frequency $\frac{P}{2N}$ that will contain energy rather than the zero frequency. In order for the harmonic number $\frac{P}{2N}$ to equal a whole number, the
period \( P \) must chosen to be an integer multiple of \( N \) such that \( P \geq 2N \). The factor of 2 in the right hand side of Equation 4.77 accounts for the fact that energy at \( \frac{\pi}{N} \) as well as \(-\frac{\pi}{N}\) will contribute to the zeroth FSF frequency.

Since the frequency \( \frac{\pi}{N} \) does not line up with the center frequencies of the FSF weighting functions, the energy at this frequency will be included not only as zero frequency energy but also in the energy corresponding to the higher FSF frequencies. This can be seen from Figure 4.17, as \( \frac{\pi}{N} \) is included in the bandwidths of \( W^0 \) and \( W^1 \). This frequency is not part of the bandwidth of \( W^2 \) but it does fall on the shoulders of this weighting function and hence it will add minutely to the energy corresponding to \( \frac{4\pi}{N} \) as well as \( \frac{6\pi}{N} \), \( \frac{8\pi}{N} \) etc (Smektala, 1998). An adjustment to Equation 2.37 is therefore also necessary for \( k = 1 \):

\[
U\left( \frac{P}{N} \right) = \sqrt{\frac{\left( \lambda_1 - 0.405\lambda_0 \right)}{M}}
\]

(4.78)

The contribution of the energy at \( \frac{\pi}{N} \) to the frequencies \( k = 2, \ldots, \frac{n-1}{2} \) is considered negligible, and therefore Equation 2.37 remains unchanged for these higher frequencies.

The final adjustment that needs to be made is in Equation 4.74 for the frequency cor-
Figure 4.17: Magnitude of FSF Weighting Functions, $W^0 (-)$, $W^1 (-)$, $W^2 (-)$

responding to $k = 0$.

$$
\lambda_0^{w_k^2} = \left| \frac{H(e^{j\pi/N})}{Q(e^{j\omega_0})} \right|^2 \left( \lambda_0^{opt} - \lambda_0^1 \frac{M_2}{M_1} \frac{\lambda_0^{f_1}}{\left| H(e^{j\pi/N}) \right|^2} \right) 
$$

(4.79)

Note that $|H(e^{j\omega_0})|^2$ has been replaced by $|H(e^{j\pi/N})|^2$. As a result, the problem relating to Equation 4.75 is avoided. It would also be technically correct to replace $|Q(e^{j\omega_0})|^2$ with $|Q(e^{j\pi/N})|^2$, but the latter cannot be estimated easily within the setup of the two-phase approach. Therefore, the estimate of $|Q(e^{j\omega_0})|^2$ will have to be used as an approximation to $|Q(e^{j\pi/N})|^2$. Equation 4.74 remains unchanged for $k = 1, \ldots, N-1$.

4.6 Summary

This chapter addressed the problem of obtaining a desired FSF spectrum in closed-loop identification experiments. The objective of the experiments was to obtain a particular distribution of energy among the FSF frequencies which minimizes the variance error in FSF-generated step response estimates. Section 4.1 introduced some of the issues that must be considered when designing an optimal experiment for FSF model identification.
An expression for the optimal reference signal was developed, but was shown to depend on unknown parameters of the closed-loop system. The unknown parameters include the weighting function, the noise model, and the background spectrum. A two-phase approach to closed-loop identification experiments was proposed.

Section 4.2 discussed the objectives of the first phase of the experiment. It was suggested that the use of a periodic reference signal allows for an efficient set of computations to be used at the end of the first phase. On-line estimation of the weighting function, noise model, and background spectrum occurs at the end of the first phase of the experiment. Due to some unique properties of the FSF model and periodic excitation, the energy present in the first phase of the experiment also contributes to the final estimation of the process model, which occurs at the end of the second phase.

In Section 4.3, the technical and practical considerations that must be made in the second phase of the experiment are analyzed. The objective of the second phase of the experiment is to ensure that the input signal prefILTERED by the inverse of the noise model, when considered as a whole, is optimal in the defined sense. Upon fixing the total desired input energy, the expression given for the optimal second-phase reference spectrum can be used for signal design. The end of the second phase marks the end of the optimal experiment, and the input-output data may then be used for off-line identification of the process and noise models.

A series of simulation studies were used in Section 4.4 to illustrate the application of the proposed two-phase approach to optimal closed-loop experiments. A step-by-step example was provided in which each aspect of the two-phase approach was discussed in detail. The results of this simulation also proved the effectiveness of the two-phase in achieving a desired FSF spectrum. A series of MonteCarlo simulations was also provided to verify the reproducibility of the two-phase approach, as well as investigating the importance of accurately modelling the weighting function. It was suggested that it may be preferable to apply a SLS algorithm to the estimation of the weighting function instead of a GLS algorithm.

Section 4.5 shows that it is not possible to obtain a desired FSF spectrum by use of the two-phase approach, even with perfect knowledge of the weighting function, noise model, and background spectrum. The presence of the transition point and cross-spectrum will in
most cases have an adverse effect on the optimality of the experiment. The section does however provide an expression for the upper bound on the difference between the desired and achieved FSF energies from a given two-phase experiment. Section 4.5 also discusses some complications relating to non-stationary noise. A simple modification to the signal design algorithm employed by the two-phase approach makes it applicable to systems corrupted by non-stationary noise.
Chapter 5

Conclusions and Recommendations

In this thesis, a new two-phase methodology to closed-loop identification was proposed. The two-phase experiment attempts to generate a data set with the optimal FSF spectrum, such that the variance of the FSF-generated step response estimate is minimized. The proposed two-phase approach can be applied to any linear closed-loop system, requiring only an \textit{a priori} estimate of the process settling time. This thesis also reviewed some practical aspects of closed-loop identification, such as the role of the noise model and data pre filters. Presented below are the conclusions of this work, followed by a series of recommendations for future consideration in the area of optimal closed-loop identification.

5.1 Conclusions

1. In Chapter 3 it was shown that the noise model plays a critical role when applying a prediction error method directly to input-output data collected from closed-loop systems. Prefiltering the input-output data with the inverse of the true noise model represents the best-case scenario for closed-loop identification. In this situation the contribution of the feedback in estimating the process model is maximized, and the process model is unbiased.

2. The use of periodic excitation allows for the application of a computationally efficient averaging routine which can be used to recover the noise properties of an open or closed-loop system. The averaging routine can be applied to systems corrupted by stationary or non-stationary noise upon appropriate pretreatment of the experimental data. The effectiveness of the averaging routine was verified by simulation studies.
3. Given an estimate of the process settling time, an input signal consisting of energy at the relatively small set of frequencies \( \frac{2\pi k}{N} \) for \( k = 0, \ldots, \frac{N-1}{2} \) can be designed for use in an FSF identification experiment. A signal consisting of a sum of sinusoids is considered to be the most efficient class of signals for FSF model identification. A multisine can be designed to contain all of its energy in the FSF frequencies. Any energy away from the FSF frequencies is attenuated by the narrow band pass nature of the FSF filters.

4. The design of an optimal reference signal for closed-loop identification requires knowledge of the weighting function, noise model, and background spectrum. Exploiting some of the properties of the FSF model and periodic excitation, these quantities can be estimated efficiently from data collected in the first phase of the proposed two-phase experiment. Using knowledge of the weighting function, noise model, and background spectrum, a reference signal can be designed and implemented in the second phase of the experiment.

5. The proposed two-phase approach was implemented in a series of simulated identification experiments. The results showed that the new methodology is reasonably effective, although the FSF spectra do not exactly match the desired distribution. It was shown that it is impossible for the FSF spectrum to exactly match the desired one, even with perfect knowledge of the weighting function, noise model, and background spectrum. The FSF energies will however be within a certain range around the desired energies.

6. The simulation studies showed that the application of a GLS algorithm to the estimation of the weighting function may degrade the overall effectiveness of the two-phase approach. Alternatively, a SLS estimator can be used without compromising consistency in the estimate of the weighting function.

5.2 Recommendations

Future work should address the following issues:

1. Constraint-handling capabilities in the multisine design algorithm. Section 4.2.4 sug-
gests using a trial-and-error application of the van den Bos algorithm for the construction of a multisine that satisfies a possible constraint on the peak amplitude. Alternatively, it may be possible to reformulate the algorithm to construct a multisine that contains the maximum possible energy while satisfying a specified bound on the amplitude of the signal.

2. Adaptation of the two-phase approach to the identification of multi-input multi-output (MIMO) systems operating in closed-loop. The problem of signal design for identification experiments becomes more complex as the number of inputs and/or outputs increases, particularly in closed-loop. The proposed two-phase approach forms a basis for work in this area.

3. Development of an efficient means for estimating the process settling time. The frequency content of the excitation signals used in the proposed two-phase approach requires an a priori estimate of the process settling time. One suggestion would be to excite the system with a pseudo random binary signal (PRBS) for a short duration prior to the first phase of the experiment. The input-output data could then be used to estimate the process step response for a given range of settling times. Using some indicator of under- or over-modelling, an estimate of the settling time could be computed on-line, allowing the two-phase methodology to be executed. The data collected in this preliminary step could also be used in the final estimation of the process step response.
Appendix A

Linking the Correlation Matrix to the Periodogram

Below is a derivation of the result presented in Equation 2.37. The expression links the diagonal elements of the FSF correlation matrix to the periodogram of the input signal.

The form of the discrete Fourier transform used here is commonly found in digital signal processing literature. That is,

\[ U(e^{j\omega_r}) = U(r) = \frac{1}{M} \sum_{t=0}^{M-1} u(t)e^{-j2\pi rt/M} \]  

(A.1)

where \( \omega_r = \frac{2\pi r}{M} \) for \( r = 0, \ldots, M - 1 \) are the DFT harmonic frequencies, and where \( M \) is the length of \( u(t) \). The periodogram is defined simply as the sequence \( |U(r)|^2 \), and the Fourier amplitude spectrum is the moduli of the Fourier coefficients, i.e. \( |U(r)| \).

An alternative expression to the one given in Equation 2.30 is also given in Wang and Cluett (1997):

\[ \lambda_k = \sum_{i=0}^{M-1} \left| F^k(e^{j\omega_i}) \right|^2 |U(e^{j\omega_i})|^2 \]  

(A.2)

where \( \omega_i = \frac{2\pi i}{M} \) and \( F^k(e^{j\omega_i}) \) is the FSF filter defined by Equation 2.9.

The weighting function, \( \left| F^k(e^{j\omega_i}) \right|^2 \), has a very narrow-bandpass nature in that it behaves like a delta function at the frequency \( \frac{2\pi k}{M} \). Properties of the weighting function include (Smektala, 1998):

1. The magnitude of the weighting function at the center frequency of the filter is equal to unity.

\[ \lim_{\omega_i \to \frac{2\pi k}{M}} \left| F^k(e^{j\omega_i}) \right|^2 = 1 \]  

(A.3)
2. At all frequencies, \(\frac{2\pi k}{N} \pm \frac{2\pi m}{N}, \left| F^k(e^{j\omega}) \right|^2 = 0\), where \(0 < m < N - 1\) is a positive integer. This means that the weighting function is zero at the center frequencies of other filters. For \(m = 1\), the distance between the nearest two zeros of \(\left| F^k(e^{j\omega}) \right|^2\) from the center frequency forms the bandwidth of the weighting function and is equal to \(\frac{2\pi}{N}\).

3. The weighting function has \(N\) local maxima. The other maxima, excluding the global one occurring at the center frequency, occur at \(\omega'_k = \frac{2\pi k}{N} \pm \left( \frac{2\pi m}{N} + \frac{\pi}{N} \right)\). The maximum value decreases as \(m\) increases, i.e. as the frequency \(\omega'_k\) moves away from the center frequency \(\frac{2\pi k}{N}\).

Assume that the energy in the input signal \(u(t)\) is only at the center frequencies of the FSF filters, i.e. \(\frac{2\pi k}{N}\) for \(k = 0, \ldots, \frac{n-1}{2}\). That is, assume the periodogram is non-zero at the frequencies \(\frac{2\pi k}{N}\) for \(k = 0, \ldots, \frac{n-1}{2}\), and zero elsewhere. From Property 1,

\[
\left| F^k(e^{j\omega_k}) \right|^2 = 1
\]  
(A.4)

Equation A.2 then reduces to

\[
\lambda_k = \sum_{t=0}^{M-1} \left| U(e^{j2\pi k}) \right|^2
\]  
(A.5)

\[
= M \left| U(e^{j2\pi k/N}) \right|^2
\]  
(A.6)

Referring back to Equation 2.30, it is important to note that these two expressions for the FSF spectrum energies differ by a factor of \(N\). The expression given in Equation 2.30 is more general, and assumes that the periodogram is relatively smooth in the vicinity of \(\omega_k\). In such a circumstance, input energy within the entire bandwidth of the weighting function, as mentioned in Property 2, will contribute to the FSF spectrum energy. This differs from the expression given in Equation A.6, where it has been assumed that all energy within the bandwidth of the weighting function is concentrated at its center frequency. This would be the case, for example, if the input was a linear combination of sinusoids at the FSF frequencies.

To continue with the derivation, the link between the FSF frequencies and the DFT harmonics must be established. Assuming we are forming a multisine input of period \(P\)
and that the process settling time $N$ is known, by setting $M = P$ in Equation A.1 we can solve the following for the DFT harmonic number $r$:

\[
\frac{2\pi r}{P} = \frac{2\pi k}{N}
\]

The DFT harmonics $r$ for $r = 0, \ldots, P - 1$ that need to be excited are therefore the set given by

\[
r = \frac{P k}{N}
\]

for $k = 0, \ldots, \frac{N-1}{2}$. The total length of the multisine is chosen as $M$; that is, it will be composed of $M/P$ periods. The desired set of FSF spectrum energies for one period becomes $\frac{\lambda_k P}{M}$, for $k = 0, \ldots, \frac{N-1}{2}$. Noting this normalization, Equation A.6 may now be expressed as

\[
\frac{\lambda_k P}{M} = P \left| U\left(\frac{P k}{N}\right)\right|^2
\]

Rearranging for the desired Fourier amplitude spectrum gives the following:

\[
U\left(\frac{P k}{N}\right) = \sqrt{\frac{\lambda_k}{M}}
\]

This ends the derivation of Equation 2.37.
Appendix B

Transformations of FSF Spectra

This appendix provides some background information on the proposition made in Equation 4.3:

$$\lambda_k \approx |S(e^{i\omega_k})|^2 |K(e^{i\omega_k})|^2 \lambda_k^{\text{rep}} + \lambda_k^{fb}$$  \hspace{1cm} (B.1)

Equation B.1 can be simplified by noting Equation 4.6 to give

$$\lambda_k \approx |Q(e^{i\omega_k})|^2 \lambda_k^{\text{rep}} + \lambda_k^{fb}$$  \hspace{1cm} (B.2)

The term $\lambda_k^{fb}$ represents the contribution of the noise to the FSF energy in the input signal when $y_{sp}(t) = 0$, and is referred to as the 'background spectrum'. The term $|Q(e^{i\omega_k})|^2 \lambda_k^{\text{rep}}$ arises from passing a deterministic signal $y_{sp}(t)$ with a FSF spectrum of $\lambda_k^{\text{rep}}$ through a stable linear filter $Q(z)$. To understand this further, recall Equation 2.30 presented here in terms of the reference signal:

$$\lambda_k^{\text{rep}} \approx \frac{M}{N} |Y_{sp}(e^{i\omega_k})|^2$$  \hspace{1cm} (B.3)

That is, the FSF spectrum $\lambda_k^{\text{rep}}$, for $k = 0, \ldots, \frac{n-1}{2}$ is proportional to the periodogram of $y_{sp}(t)$. Therefore, this analysis of the transformation of FSF spectra will be done in terms of the associated periodograms.

Let $u(t)$ and $y_{sp}(t)$ be related by the strictly stable system $Q(z)$ as follows:

$$u(t) = Q(z)y_{sp}(t)$$  \hspace{1cm} (B.4)

The signal $y_{sp}(t)$ obeys $|y_{sp}(t)| \leq C_W$ for all $t$. Then from Theorem 2.1 in Ljung (1999),

$$U(e^{i\omega}) = Q(e^{i\omega})Y_{sp}(e^{i\omega}) + R(e^{i\omega})$$  \hspace{1cm} (B.5)
where

\[ |R(e^{j\omega})| \leq \frac{2C_W C_G}{N} \]  \hspace{1cm} (B.6)

with

\[ C_G = \sum_{k=1}^{\infty} k |q_k| \]  \hspace{1cm} (B.7)

The coefficients \( b_k \) are simply the impulse response coefficients of \( Q(z) \). Ljung (1999) notes that if \( y(t) \) is periodic with period \( N \), then \( R(e^{j\omega}) \) in Equation B.5 is zero for \( \omega = \frac{2\pi k}{N} \). In this case, the following is true:

\[ |U(e^{j\omega})|^2 \cdot |Q(e^{j\omega})|^2 \cdot |Y_{sp}(e^{j\omega})|^2 \]  \hspace{1cm} (B.8)

Otherwise, Equation B.8 can serve as an approximation when \( y(t) \) is aperiodic. Multiplying both sides by \( \frac{M}{N} \) and noting Equations 2.30 and B.3 gives

\[ \lambda_k \approx \left| Q(e^{j\omega_k}) \right|^2 \lambda_{sp} \]  \hspace{1cm} (B.9)

Assuming the effect of the background spectrum \( \lambda_k^{fb} \) is additive, Equation B.9 can be modified as follows to give the desired result:

\[ \lambda_k \approx \left| Q(e^{j\omega_k}) \right|^2 \lambda_{sp} + \lambda_k^{fb} \]  \hspace{1cm} (B.10)

The discussion presented in Section 4.5.2 refutes the assumption that the background spectrum is additive in Equation B.10. However, this expression is still used as a tool for the optimal experiment designs presented in this thesis.
Appendix C

Magnitude of the Cross-Spectrum

Below is a derivation of the result presented in Equation 4.60. The expression provides an upper bound on the magnitude of the cross-spectrum. For any two uncorrelated signals $u_1(t)$ and $u_2(t)$, the upper bound on the cross-spectrum can be described as follows:

$$\left| \sum_{t=0}^{M-1} [\phi_1^k(t)\phi_2^k(t) + \phi_2^k(t)\phi_1^k(t)] \right| \leq 2\sqrt{\lambda_{k,1}\lambda_{k,2}} \quad (C.1)$$

where

$$\lambda_{k,1} = \sum_{t=0}^{M-1} \phi_1^k(t)\phi_1^k(t) \quad (C.2)$$

and similarly for $\lambda_{k,2}$. Let

$$\phi_1(t) = a_1(t) + b_1(t)j \quad (C.3)$$

and

$$\phi_2(t) = a_2(t) + b_2(t)j \quad (C.4)$$

It is easy to show that the FSF energies can be expressed as

$$\lambda_{k,1} = \sum_{t=0}^{M-1} [a_1^2(t) + b_1^2(t)] \quad (C.5)$$

$$\lambda_{k,2} = \sum_{t=0}^{M-1} [a_2^2(t) + b_2^2(t)] \quad (C.6)$$

Similarly, the cross spectrum can be expressed as

$$\sum_{t=0}^{M-1} [\phi_1^k(t)\phi_2^k(t) + \phi_2^k(t)\phi_1^k(t)] = 2 \sum_{t=0}^{M-1} [a_1(t)a_2(t) + b_1(t)b_2(t)] \quad (C.7)$$

Substituting Equations C.5, C.6, and C.7 into Equation C.1 gives

$$2 \left| \sum_{t=0}^{M-1} [a_1(t)a_2(t) + b_1(t)b_2(t)] \right| \leq 2 \sqrt{\left( \sum_{t=0}^{M-1} [a_1^2(t) + b_1^2(t)] \right) \left( \sum_{t=0}^{M-1} [a_2^2(t) + b_2^2(t)] \right)} \quad (C.8)$$
The proof of Equation C.8 is related to the Cauchy-Schwarz inequality. Define the vectors \( \mathbf{x}_1 \) and \( \mathbf{x}_2 \) as follows:

\[
\mathbf{x}_1 = [a_1(0), a_1(1), \ldots, a_1(M - 1), b_1(0), b_1(1), \ldots, b_1(M - 1)] \\
\mathbf{x}_2 = [a_2(0), a_2(1), \ldots, a_2(M - 1), b_2(0), b_2(1), \ldots, b_2(M - 1)]
\]  

(C.9) \hspace{2cm} (C.10)

The Cauchy-Schwarz inequality then states the following:

\[
|\mathbf{x}_1 \cdot \mathbf{x}_2| \leq ||\mathbf{x}_1|| ||\mathbf{x}_2||
\]  

(C.11)

Alternatively, Equation C.11 can be expressed in terms of the vector components:

\[
\left| \sum_{i=0}^{M-1} [a_1(t) a_2(t) + b_1(t) b_2(t)] \right| \leq \sqrt{ \left( \sum_{i=0}^{M-1} [a_1^2(t) + b_1^2(t)] \right)} \sqrt{ \left( \sum_{i=0}^{M-1} [a_2^2(t) + b_2^2(t)] \right)}
\]

(C.12)

The similarity to Equation C.8 is obvious, thus ending the proof.


