Time-interleaved analog-to-digital converter mismatch compensation using infinite impulse response filters

Xu, Saihua

2016

Xu, S. (2016). Time-interleaved analog-to-digital converter mismatch compensation using infinite impulse response filters. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/66397>

<https://doi.org/10.32657/10356/66397>

**TIME-INTERLEAVED ANALOG-TO-DIGITAL
CONVERTER MISMATCH COMPENSATION
USING INFINITE IMPULSE RESPONSE FILTERS**

XU SAIHUA

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2016

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Lim Yong Ching, for accepting me as his PhD student at the first place. I sincerely appreciate his excellent guidance, valuable suggestions and continuous support throughout the entire course of my research work. His encouragement and confidence in me helped me get through many difficult situations and gave me the strength to complete this challenging task. I could not have imagined having a more caring, understanding and patient supervisor and mentor.

Secondly, I would like to thank my fellow colleagues at Temasek Laboratories@NTU, Dr. Lee Jun Wei, Mr. Leong Jenn Gan, Dr. Yang Chunzhu, and Mr. Zhu Zhidong for generously sharing their knowledge and experience with me and their friendship. I would always cherish the great times we had all these years.

Doing a PhD has been a rewarding, and at the same time struggling experience for me. I am deeply grateful for my beloved husband, Mr. Guan Xiaochong, who was always there to share my joy when I made every single achievement and to cheer me up when things weren't going smoothly.

Lastly, I wish to express my deepest gratefulness to my parents, my sister and my mother-in-law for their relentless love, understanding and moral support throughout all these years.

Table of Contents

Acknowledgements	ii
Table of Contents.....	iii
List of Figures	vi
List of Tables.....	x
Abbreviations.....	xii
Summary	xiii
CHAPTER 1 : INTRODUCTION	1
1.1 Motivation.....	1
1.2 Objective	3
1.3 Major Contributions of the Thesis	3
1.4 Organization of the Thesis	4
CHAPTER 2 : LITERATURE REVIEW	7
2.1 TIADC Architecture	7
2.2 Sources of Mismatch Errors and Their Effects	9
2.2.1 Sources of Mismatches.....	10
2.2.2 Effects of the Mismatches in the Frequency Domain.....	13
2.3 Mismatch Estimation and Calibration.....	17
2.3.1 Offset Mismatch.....	18
2.3.2 Gain Mismatch	20
2.3.3 Timing Mismatch	22
2.3.4 Frequency Response Mismatch.....	33
2.4 Chapter Summary	36
CHAPTER 3 : DESIGN PRINCIPLE OF THE IIR FILTERS FOR TIADC MISMATCH COMPENSATION.....	37
3.1 Introduction.....	37
3.2 The Proposed IIR Mismatch Compensation Filter.....	37
3.3 WLS Optimization Approach	46

3.4	Discussion	48
3.5	Chapter Summary	52
CHAPTER 4 : STABILITY CRITERION OF THE PROPOSED IIR MISMATCH COMPENSATION FILTER		53
4.1	Introduction	53
4.2	Equivalent MIMO Representation of a PTV System	54
4.2.1	MIMO Representation of an <i>M</i> -PTV IIR Filter	54
4.2.2	MIMO Representation of an <i>M</i> -PTV FIR Filter	60
4.3	Impulse Response of an <i>M</i> -PTV IIR Filter	62
4.4	Alternative Approach to Derive the MIMO Transfer Function Matrix	65
4.5	Chapter Summary	70
CHAPTER 5 : METHODS TO ENSURE THE STABILITY OF THE <i>M</i>-PTV IIR COMPENSATION FILTER		72
5.1	Introduction	72
5.2	<i>Method-1</i> : By Increasing the Integer Group Delay	73
5.3	<i>Method-2</i> : By Constraining the Magnitudes of the Denominator Coefficients	79
5.4	Minimum Group Delay Required to Achieve Stability	82
5.4.1	Timing Mismatch Case	82
5.4.2	Bandwidth Mismatch Case	93
5.5	Chapter Summary	102
CHAPTER 6 : SIMULATION EXAMPLES AND DISCUSSION		103
6.1	Introduction	103
6.2	Timing and Bandwidth Mismatch Example	103
6.3	Bandwidth Mismatch Example	114
6.4	Timing Mismatch Example	121
6.5	Performance in the Presence of Identification Errors	127
6.6	Filter Redesign When Mismatch Parameters Drift	129
6.7	Pipeline Implementation of the IIR Compensation Filter	136
6.8	Chapter Summary	140
CHAPTER 7 : FINITE WORDLENGTH EFFECTS IN THE <i>M</i>-PTV IIR FILTER		141
7.1	Introduction	141
7.2	Roundoff Noise Introduced by Input Signal Quantization	142
7.2.1	Quantization Noise in an Ideal ADC	142
7.2.2	Quantization Noise in a Non-ideal TIADC	143

7.3	Roundoff Noise Introduced by Multiplier Output Truncation	150
7.4	Coefficient Quantization	156
7.5	Limit Cycle	161
7.6	Chapter Summary	163
CHAPTER 8 : CONCLUSION AND FUTURE WORK		165
8.1	Conclusion	165
8.2	Future Work	168
Author's Publications		169
Bibliography		170
Appendix A: Detailed Derivation of the Equivalent MIMO System Using State Space Representation		184
Appendix B: Eigenvalues of the Matrices $\tilde{A}$ and $\tilde{A}'$		187

List of Figures

Figure 2.1	An ideal ADC is modelled by a uniform sampler and a quantizer.	7
Figure 2.2	Spectrum of (a) the bandlimited analog signal $x_a(t)$ and (b) the ideally sampled sequence $\tilde{x}[n]$	8
Figure 2.3	Architecture of an M -channel TIADC and the M clock signals driving the sub-ADCs. 9	
Figure 2.4	Transfer characteristics of different ADCs experiencing different gains.	10
Figure 2.5	An example of a two-channel TIADC. (a) Ideal sampling instants (dashed line). (b) Real sampling instants (solid line).	11
Figure 2.6	An equivalent model of the m -th sub-ADC channel.	13
Figure 2.7	Output spectrum of a four-channel TIADC over $[0, \Omega_s]$ assuming a sinusoidal input with frequency Ω_{in} . (a) Gain and offset mismatches. (b) Timing and frequency response mismatches.	15
Figure 2.8	The LMS loop for analog background tracking and correction of the offset and gain mismatches in [17].	18
Figure 2.9	The LMS loop for digital background tracking and correction of the gain mismatch in [19].	21
Figure 2.10	A two-rank S/H system to avoid timing mismatch caused by nonideal sampling instants.	22
Figure 2.11	(a) An all digital multirate filter bank structure for the reconstruction of uniform samples from the nonuniformly sampled TIADC output. (b) An equivalent hybrid model with continuous-time analysis bank and discrete-time synthesis bank.	24
Figure 3.1	A hybrid analog-digital system modelling an M -channel TIADC.	38
Figure 3.2	An example of a three-channel TIADC demonstrating the relationships between the ideal ADC output sample $\tilde{x}[n]$, the sub-ADC output sample $x_m[k]$, the compensated sub-ADC output sample $\hat{x}_m[k]$, the nonideal TIADC output sample $x[n]$, and the compensated value for $x[n-3]$, $y[n]$	40
Figure 3.3	Multichannel IIR filter structure implementation for computing $\hat{x}_1[k]$	42
Figure 3.4	The general multichannel IIR filter structure with M channels for estimating the low rate sequence $\hat{x}_m[k]$	42
Figure 3.5	(a) A general time-varying filter. (b) Direct form I implementation of the proposed time-varying IIR filter for TIADC mismatch compensation.	44
Figure 4.1	Direct form I implementation of the proposed M -PTV IIR compensation filter with the state variables indicated.	54
Figure 4.2	Equivalent MIMO system with vector input and vector output of the M -PTV IIR filter.	59
Figure 4.3	Equivalent filter bank realization of the M -PTV IIR filter.	65
Figure 4.4	Equivalent high rate MIMO system of the M -PTV IIR filter.	68
Figure 5.1	Effect of the group delay η on the WLS error.	76

Figure 5.2	(a) Effect of the group delay η on the maximum pole radius p_{max} . (b) Zoomed in view where the range of η is from 6 to 14.	77
Figure 5.3	Effect of the denominator coefficient weight on the WLS error.	81
Figure 5.4	Effect of the denominator coefficient weight on the maximum pole radius p_{max}	81
Figure 5.5	Occurrence of k in the 5670 different sets of design parameters where k is the range of consecutive numbers that more than 99% of the 10 thousand cases of η_{min} fall on for each set of design parameters. Timing mismatch case.	83
Figure 5.6	Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and ε_{max} , M is fixed to 2. Timing mismatch case.	85
Figure 5.7	Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and M , ε_{max} is fixed to 0.01. Timing mismatch case.	85
Figure 5.8	Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and ε_{max} , M is fixed to 4. Timing mismatch case.	86
Figure 5.9	Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and M , ε_{max} is fixed to 0.05. Timing mismatch case.	86
Figure 5.10	Plots of $\bar{\eta}_{min}$ versus ε_{max} for different values of N_a . Timing mismatch case. (a) The case where $N_b = 3$ and $M = 2$. (b) The case where $N_b = 9$ and $M = 5$	87
Figure 5.11	Occurrence of k in the 3402 different sets of design parameters where k is the range of consecutive numbers that more than 99% of the 10 thousand cases of η_{min} fall on for each set of design parameters. Bandwidth mismatch case.	94
Figure 5.12	Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and ε_{max} , M is fixed to 5. Bandwidth mismatch case.	95
Figure 5.13	Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and M , ε_{max} is fixed to 0.3. Bandwidth mismatch case.	95
Figure 5.14	Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and ε_{max} , M is fixed to 4. Bandwidth mismatch case.	96
Figure 5.15	Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and M , ε_{max} is fixed to 0.2. Bandwidth mismatch case.	96
Figure 5.16	Plots of $\bar{\eta}_{min}$ versus ε_{max} for different values of N_a . Bandwidth mismatch case. (a) The case where $N_b = 3$ and $M = 2$. (b) The case where $N_b = 10$ and $M = 5$	97
Figure 6.1	Simulated frequency spectrums in Example 6.1. (a) Input. (b) Uncompensated output. (c) Compensated output using FIR filter. (d) Compensated output using IIR-1 filter with $p_{max} = 0.932$. (e) Compensated output using IIR-2 filter with $p_{max} = 0.970$. (f) Compensated output using IIR-2 filter with $p_{max} = 0.930$. The total number of coefficients for (c) through (f) is each $4 \times 30 = 120$	106
Figure 6.2	Magnitude responses of $H_m(e^{j\omega})$, $m = 0, 1, \dots, M-1$ for the filters designed in Example 6.1. (a) FIR filter. (b) IIR-1 filter with $p_{max} = 0.932$. (c) IIR-2 filter with $p_{max} = 0.970$. (d) IIR-2 filter with $p_{max} = 0.930$	112
Figure 6.3	Magnitude responses of the LTI IIR filters, $R_m(e^{j\omega})$, $m = 0, 1, \dots, M-1$ in the equivalent filter bank representation for the filters designed in Example 6.1. (a) FIR	

	filter. (b) IIR-1 filter with $p_{max} = 0.932$. (c) IIR-2 filter with $p_{max} = 0.970$. (d) IIR-2 filter with $p_{max} = 0.930$	113
Figure 6.4	Simulated frequency spectrum in Example 6.2, assuming multi-tone input. (a) Input. (b) Uncompensated output. (c) Compensated output using FIR filter. (d) Compensated output using IIR-1 filter with $p_{max} = 0.896$. (e) Compensated output using IIR-2 filter with $p_{max} = 0.950$	117
Figure 6.5	Simulated frequency spectrums in Example 6.2, assuming bandlimited Gaussian noise input. (a) Uncompensated output. (b) Compensated output using the FIR filter. (c) Compensated output using IIR-1 filter with $p_{max} = 0.896$. (d) Compensated output using IIR-2 filter with $p_{max} = 0.950$	118
Figure 6.6	Simulated frequency spectrums for $\beta = 1$ in Example 6.3. (a) Uncompensated output. (b) Compensated output using FIR filter; 4×47 coefficients are required for 100 dB SFDR. (c) Compensated output using IIR filter; only 4×33 coefficients are required and the SFDR is 103.2 dB.	122
Figure 6.7	Worst case SFDR in the presence of identification error for the FIR and IIR filters designed in Example 6.2.	128
Figure 6.8	Worst case SFDR in the presence of identification error for the FIR and IIR filters designed in Example 6.3.	129
Figure 6.9	Maximum and minimum values of p_{max} of the IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the IIR filters in Fig. 6.1(d) and (f) versus r_{dr}	131
Figure 6.10	Maximum and minimum values of SFDR of the FIR, IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the filters in Fig 6.1(c), (d) and (f) versus r_{dr}	132
Figure 6.11	Maximum and minimum values of p_{max} of the IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the IIR filters in Fig. 6.4(d) and (e) versus r_{dr}	133
Figure 6.12	Maximum and minimum values of SFDR of the FIR, IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the filters in Fig 6.4(c), (d) and (e) versus r_{dr}	133
Figure 6.13	Maximum and minimum values of p_{max} of the IIR-1 filters redesigned using the same design parameters as those in Table. 6.9 versus r_{dr} for each value of β	135
Figure 6.14	Maximum and minimum values of SFDR of the FIR and IIR-1 filters redesigned using the same design parameters as those in Table. 6.9 versus r_{dr} for each value of β	136
Figure 7.1	(a) An ideal ADC modelled by a sampler and a quantizer. (b) An equivalent quantization noise model for the ADC.	142
Figure 7.2	(a) Quantization noise model for a TIADC. (b) Quantization noise $e_x[n]$ propagates through the M -PTV IIR compensation filter $H_n(z)$ and appears at the output as $e_y[n]$	144
Figure 7.3	Quantization noise model for the equivalent MIMO system of the M -PTV IIR filter.	145
Figure 7.4	Noise amplification in each channel of the IIR compensation filter with respect to the maximum pole radius p_{max}	149

Figure 7.5	Noise model for multiplier output truncation in an M -PTV IIR filter. (a) The errors due to multiplier output truncation are modelled by the addition of the noise sources $e_{a,p}[n]$, $p=0, 1, \dots, N_a-1$ and $e_{b,q}[n]$, $q=0, 1, \dots, N_b$. (b) The multiplier output truncation noise sources are replaced by a single noise source $e_c[n]$	151
Figure 7.6	Simulated output frequency spectrum of the IIR filters. Stability is achieved using <i>Method-1</i> . (a) Infinite precision and $\eta = 20$. (b) Infinite precision and $\eta = 21$. (c) 18-bit coefficient wordlength and $\eta = 20$. (d) 18-bit coefficient wordlength and $\eta = 21$	158
Figure 7.7	$\sigma_{spurious}^2$ of the FIR filter and the IIR filters with $\eta = 20$ and $\eta = 21$ versus coefficient wordlength.....	159
Figure 7.8	Simulated output frequency spectrum of the IIR filters. Stability is achieved using <i>Method-2</i> . (a) Infinite precision with $p_{max} = 0.970$. (b) Infinite precision with $p_{max} = 0.900$. (c) 16-bit coefficient wordlength with $p_{max} = 0.970$. (d) 16-bit coefficient wordlength with $p_{max} = 0.900$	160
Figure 7.9	$\sigma_{spurious}^2$ of the FIR filter and the IIR filters with $p_{max} = 0.970$ and $p_{max} = 0.900$ versus coefficient wordlength.....	161
Figure 7.10	A first order IIR filter with a quantizer in the feedback loop.	162

List of Tables

Table 3.1	List of the notations for the various signals appearing in Chapter 3	39
Table 4.1	Filter coefficients of the PTV IIR filter in Example 4.1.....	70
Table 5.1	Mismatch parameters for the 4-Channel TIADC in Example 5.1	76
Table 5.2	Mismatch parameters for the 4-channel TIADC in Example 5.2.....	80
Table 5.3	List of the values of the design parameters simulated to identify η_{min} . Timing mismatch case.	83
Table 5.4	The coefficients in (5.16) for approximating $\bar{\eta}_{min}$. Timing mismatch case.	89
Table 5.5	All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 2$. Timing mismatch case.	90
Table 5.6	All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 4$. Timing mismatch case.	91
Table 5.7	All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 5$. Timing mismatch case.	92
Table 5.8	List of the values of the design parameters simulated to identify η_{min} . Bandwidth mismatch case.	93
Table 5.9	The coefficients in (5.19) for approximate $\bar{\eta}_{min}$. Bandwidth mismatch case.....	99
Table 5.10	All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 2$. Bandwidth mismatch case.	99
Table 5.11	All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 4$. Bandwidth mismatch case.	100
Table 5.12	All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 5$. Bandwidth mismatch case.	101
Table 6.1	Mismatch parameters for the 5-channel TIADC in Example 6.1	104
Table 6.2	Coefficient values for the FIR compensation filter in Example 6.1	107
Table 6.3	Coefficient values for the IIR-1 filter with $p_{max} = 0.932$ in Example 6.1	108
Table 6.4	Coefficient values for the IIR-2 filter with $p_{max} = 0.970$ in Example 6.1	109
Table 6.5	Coefficient values for the IIR-2 filter with $p_{max} = 0.930$ in Example 6.1	110
Table 6.6	Bandwidth mismatch parameters for the 5-channel TIADC in Example 6.2	114
Table 6.7	Comparison of the implementation complexity and performance of the FIR and IIR filters in Example 6.2.	119
Table 6.8	Timing mismatch parameters for the 5-channel TIADC in Example 6.3.....	121
Table 6.9	Coefficient savings in percentage for various values of β	123
Table 6.10	Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the FIR filters in Table 6.9 where the filters are designed with fixed SFDR	124

Table 6.11	Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filters in Table 6.9 where the filters are designed with fixed SFDR	124
Table 6.12	Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the FIR filters whose design parameters are fixed to $N_a = 47$, $N_b = 0$ and $\eta = 23$	125
Table 6.13	Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filters whose design parameters are fixed to $N_a = 30$, $N_b = 4$ and $\eta = 22$	126
Table 6.14	Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filters with $N_a = 28$, $N_b = 6$, $\eta = 22$ and a uniform weight of 1.789×10^{-11}	127
Table 6.15	Coefficient values of the equivalent filter bank for IIR-1 filter designed in Example 6.2.	138
Table 7.1	Noise amplifications at the output ports of the filters designed in Example 6.1. The values are obtained through analysis.	147
Table 7.2	Noise variances at the output ports of the filters designed in Example 6.1. Assume 10-bit quantization noise input.	148
Table 7.3	Noise amplifications at the output ports of the filters designed in Example 6.1. The values are obtained through simulations.	148
Table 7.4	Noise amplifications at the output ports of the filters designed in Example 6.2. The values are obtained through analysis.	150
Table 7.5	Noise amplifications at the output port of the filters designed in Example 6.2. The values are obtained through simulations.	150
Table 7.6	The energy of each element of the transfer function matrix $\mathbf{H}^{deno}(z)$ for the IIR filters designed in Example 6.1.	154
Table 7.7	The noise variances due to multiplier output wordlength truncation for the FIR and IIR filters designed in Example 6.1. The results are obtained through analysis.	155
Table 7.8	The noise variances due to multiplier output wordlength truncation for the FIR and IIR filters designed in Example 6.1. The results are obtained through simulation.	155
Table 7.9	The impulse response of the first order IIR filter shown in Figure 7.10.	162

Abbreviations

ADC	Analog-to-digital converter
DAC	Digital-to-analog converter
DFT	Discrete Fourier transform
DSP	Digital signal processing or digital signal processor
DTFT	Discrete-time Fourier transform
FFT	Fast Fourier transform
FIR	Finite impulse response
FPGA	Field-programmable gate array
GS/s	Giga samples per second
GMAC/s	Giga multiply accumulate operations per second
IC	Integrated circuit
IDFT	Inverse discrete Fourier transform
IF	Intermediate frequency
IFFT	Inverse Fast Fourier transform
<i>i.i.d.</i>	Independent, identically distributed
IIR	Infinite impulse response
LMS	Least mean squares
LTi	Linear time invariant
LUT	Look up table
MIMO	Multi-input multi-output
MSE	Mean squared error
PR	Perfect reconstruction
PTV	Periodic time-varying
rad/s	Radians per second
RF	Radio frequency
RMS	Root-mean-square
S/H	Sample-and-hold
SFDR	Spurious free dynamic range
SINAD	Signal to noise and distortion ratio
TIADC	Time-interleaved analog-to-digital converter
WGN	White Gaussian noise
WLS	Weighted least squares
WSCS	Wide sense cyclostationary
WSS	Wide sense stationary

Summary

Time interleaving is an effective method to achieve a high sampling rate by using a bank of low sampling rate analog-to-digital converters (ADCs). The resulting structure is a time-interleaved ADC (TIADC). If the slow ADCs in the TIADC structure are identical and the clock signals driving them are perfectly aligned, the effective resolution of the TIADC will be the same as that of the sub-ADCs'. However, this is never the case in practice. Mismatches among different sub-ADC channels introduce unwanted distortion images in the output spectrum of the TIADC, leading to degraded system resolution. Different sub-ADCs may have different constants of proportionality between input and output signals leading to gain mismatch and different DC offsets leading to offset mismatch. The sampling instants of the sub-ADCs may not be perfectly aligned leading to timing mismatch. Moreover, the front-end circuitry of different sub-ADC channels may experience different frequency characteristics leading to frequency response mismatch.

In this thesis, we first present an overview on the architecture of the TIADC. Four types of mismatch errors exist in a TIADC structure, namely, the offset, gain, timing and frequency response mismatches are then reviewed. The effects of the mismatch errors to the TIADC output in both the time and frequency domains are derived. Existing methods for correcting the mismatch errors and reconstructing the TIADC output are surveyed. Among the various correction methods, post-processing the TIADC output data in the digital domain to interpolate the ideal samples has been drawing tremendous research attention recently. Traditionally, finite impulse response (FIR) filters were used in various reconstruction structures to compensate for the TIADC mismatch error. The advantages of using FIR filters are that the FIR filters are inherently stable and that the design techniques are readily available. In this work, we introduce an unprecedented design method which utilizes infinite impulse response (IIR) filters to compensate for the mismatch errors in the TIADC. Our IIR mismatch compensation system may be interpreted as multichannel IIR filters with M input ports or an M -periodic time-varying (M -PTV) IIR filter, where M is the number of interleaving

channels in the TIADC. We demonstrate the design principle of the IIR compensation filter and introduce a weighted least squares (WLS) optimization method to obtain the filter coefficients.

In our IIR compensation system, feedback from the past estimated output is an essential part. Therefore, the stability issue of the IIR compensation filter needs to be investigated. By using the state space representation and vector processing, the IIR compensation filter is transformed into an equivalent multi-input multi-output (MIMO) system operating at one- M th of the TIADC output rate. The proposed IIR compensation filter is stable if all the eigenvalues of the *state transition matrix* of the MIMO system have magnitudes less than one. Incorporating the eigenvalue based stability criterion into the WLS optimization is difficult and would increase the design complexity. Therefore we introduce two “computer aided” design approaches for gradually achieving a stable IIR filter design; one is by increasing the integer group delay and the other is by constraining the magnitudes of the feedback coefficients of the IIR filter. Explicit proof of the first method is presented. Moreover, the relationship between the minimum group delay required to achieve stability and various design parameters are investigated through numerous simulations. The derived expressions serve as useful guidelines for selecting a proper value of group delay in the design of the IIR compensation filter.

By using simulation examples, we show that IIR filters with guaranteed stability for TIADC mismatch compensation can be designed and our IIR filters have much lower implementation complexity than the FIR counterparts when they are designed to achieve the same level of compensation performance. Specifically, for a given spurious frequency power, our IIR compensation filters yield about 30% saving in the total number of multiplier coefficients when compared to the FIR counterparts. The number of delay elements required and the number of arithmetic operations per sample in the case of the IIR compensation filters are also lower than those in the case of the FIR filters. Other important aspects such as the performance in the presence of identification error, filter redesign to cope with mismatch drifting, pipeline implementation etc. are also studied. Lastly, we discuss the finite wordlength effects in our IIR compensation filter including roundoff noise propagation, coefficient quantization and limit cycle.

CHAPTER 1 : INTRODUCTION

1.1 Motivation

Digital signal processing (DSP) finds its application in a wide range of advanced systems such as telecommunication, sonar and radar processing, sensor technology etc. Over the past decades, the rapid evolution of digital integrated circuit (IC) technology together with parallel computing architecture has made DSP systems faster and less expensive. On the other hand, real world signals need to be transformed into discrete, binary coded samples before being processed by the sophisticated software algorithms. Therefore, the availability of high speed and high resolution analog-to-digital converters (ADCs) is essential to the success of these advanced systems. In turn, the ever-developing of these advanced systems has been consistently pushing the boundaries of ADCs. For a single-core ADC, flash ADCs have the highest speed at the expense of large area, high power dissipation and limited resolution. Also, it has been pointed out in [1] that for ADCs operating at multi-Giga samples per second (multi-GS/s) rates, the effective resolution is limited by the comparator ambiguity which is related to the speed of the device technology used to fabricate the ADC. Downscaling of the device technology leads to speed improvement but makes circuit design difficult.

In order to achieve high conversion rate and good resolution while maintaining reasonable increase in the area and power dissipation within a given device technology, a time-interleaved ADC (TIADC) scheme has been proposed by Black and Hodge [2] and widely used since then. A TIADC consists of several slow sub-ADCs sampling at interleaved time instants as if they were effectively a single ADC operating at a much higher speed. TIADC has relaxed the stringent requirement on the device technology and its application can be found in many high end systems. One example would be the

radar system. The radar identifies the position and movement of an object by transmitting radio signals and analyzing the reflections. The received radio frequency (RF) signals have very high frequencies. In order for the back-end digital processor to process the received signal, the RF signal needs to go through a local mixer stage to be converted to intermediate frequency (IF) signal so that it can be sampled by an ADC. The availability of the high speed TIADC allows direct sampling of the radio signals at RF frequencies, eliminating the need for analog down conversion stages. In spite of the significant speed advantage, TIADCs have a major drawback that any mismatch between the sub-ADC channels generates undesirable distortions in the output spectrum and significantly degrades the performance of the TIADC such as the spurious free dynamic range (SFDR) and the signal to noise and distortion ratio (SINAD) [3].

Over the years, enormous research efforts have been devoted to correcting the mismatch errors and improving the TIADC performance. Among the various mismatch correction techniques proposed, digital post-processing has drawn tremendous research interest as it offers superior accuracy, great reliability, and flexibility to port to other device technologies. Traditionally, finite impulse response (FIR) filters were employed in various digital compensation structures to improve the TIADC performance. FIR filters have the advantage that they are inherently stable but may need a large number of taps to obtain the desired response.

To achieve real-time data processing, digital computation algorithms are usually implemented on field-programmable gate array (FPGA) or digital signal processor (DSP) chips. Specifically, the digital filters are implemented using the embedded multipliers and adders of the FPGAs and DSPs. The amount of hardware resources to implement a filter and the number of arithmetic operations to produce one output sample is proportional to the filter length. Further, online filter design is sometimes required. The amount of computations and hence the resource needed for online design is also determined by the filter complexity. Therefore reducing the filter complexity, while maintaining the desired compensation performance, is essential to the design of real-time digital compensation systems.

Infinite impulse response (IIR) filters, on the other hand, are known for being able to achieve sharp transition with small number of filter coefficients. Therefore, lower filter complexity is expected when

IIR filters are designed instead of FIR filters to achieve the same level of mismatch compensation performance. In this thesis, we shall present a novel TIADC compensation structure consisting of IIR filters, which compared to the traditional FIR filters, has reduced complexity for the same level of compensation performance.

1.2 Objective

The main objective of this work is to design IIR filters with guaranteed stability to compensate for the mismatch errors in a TIADC, specifically the timing and frequency response mismatches; to search for methods with low computational effort for achieving a stable IIR filter design; and to investigate the performance of the IIR compensation filters in various aspects and compare them to that of the FIR filters.

1.3 Major Contributions of the Thesis

Under the scope of work mentioned in the last section, the major contribution of this thesis can be summarized as follows:

1. A new TIADC compensation structure consisting of recursive filters was presented. The design principle of the IIR compensation filter and a weighted least squares (WLS) optimization approach to obtain the filter coefficients were demonstrated.
2. The proposed IIR compensation filter can be represented by an equivalent multi-input multi-output (MIMO) system and is stable if all the eigenvalues of the state transition matrix of the MIMO system have magnitudes less than one. However, it is difficult to incorporate the eigenvalue based stability criterion into the WLS optimization process. Therefore, two simple computer-aided design methods for gradually achieving a stable IIR filter design were introduced. One is by increasing the integer group delay and the other is by reducing the magnitudes of the

- feedback coefficients. Theoretical proof of the first method by using *Rouché's Theorem* was presented.
3. The relationship between the minimum value of group delay to achieve stability and the various design parameters, including the number of interleaving channels, the number of numerator and denominator coefficients and the scale of the timing and bandwidth mismatch errors, was investigated. The results provide good design guidelines for selecting the group delay value which would result in a stable IIR filter under various combinations of the design parameters.
 4. The performance of the IIR compensation filter was characterized and compared to that of the FIR counterpart. It was demonstrated that for the same level of mismatch compensation performance, the total number of multipliers required by our IIR filters is 30% to 35% less than that required by the FIR counterparts. Other practical aspects of the IIR compensation filter including the performance in the presence of identification error, filter redesign to cope with mismatch drift, numerical robustness, pipeline implementation, and finite wordlength effects etc. were investigated.

1.4 Organization of the Thesis

Chapter 1 presents the background and motivation of the current research work, research objective, major contribution and thesis organization.

Chapter 2 gives an overview of the TIADC architecture and the four types of mismatch errors arising from the parallel structure, namely, the offset, gain, timing and frequency response mismatches. Then the techniques proposed by various researchers in over twenty years for estimating and compensating each of the four types of mismatch errors are surveyed.

Chapter 3 introduces an unprecedented design method which utilizes IIR filters to compensate for the mismatch errors in the TIADC. The proposed IIR mismatch compensation filter can be interpreted as either multichannel IIR filters or a periodic time-varying (PTV) IIR filter. The design principle of the

IIR compensation filter is demonstrated and a WLS optimization approach to obtain the filter coefficients is presented.

Chapter 4 derives the stability criterion for our IIR mismatch compensation filter analytically. By using the state space representation and vector processing, the IIR compensation filter is transformed into an equivalent MIMO system operating at the channel sampling rate. The proposed IIR filter is stable if all the eigenvalues of the state transition matrix of the MIMO system have magnitudes less than one. Although the stability criterion is derived explicitly, it is difficult to incorporate the eigenvalue based constraint into the WLS optimization process without significantly increasing the design complexity.

Chapters 5 presents two simple methods for gradually achieving a stable IIR compensation filter design, one is by increasing the integer group delay and the other is by constraining the magnitudes of the feedback coefficients. The validity of the first method is proved theoretically by using the *Rouché's Theorem*. It is also shown by a simulation example that as the value of the group delay increases, the maximum pole radius of the IIR compensation filter decreases. The relationship between the minimum group delay required to reach stability and various design parameters, including the number of numerator and denominator coefficients, the number of interleaving channels, the scales of the timing and bandwidth mismatches, are then investigated. The results serve as useful guidelines for selecting a suitable value of group delay to achieve filter stability in the design process.

Chapter 6 contains the simulation results. By using some simulation examples, we show that (1) IIR filters with guaranteed stability for TIADC mismatch compensation can be designed and (2) compared to the FIR counterparts, our IIR filters achieve much better compensation performance when they are designed with the same total number of filter coefficients or have much lower implementation complexity when they are designed to achieve the same level of compensation performance. Specifically, for a given SFDR, our IIR compensation filters yield about 30% to 35% saving in the total number of multiplier coefficients when compared to the FIR filters. The number of delay elements required and the number of arithmetic operations per sample in the case of the IIR compensation filter are also about 30% to 35% lower than those in the case of the FIR filters. Other important aspects of the IIR compensation filter such as the performance in the presence of

identification error, filter redesign to cope with mismatch drifting, pipeline implementation etc. are also investigated.

Chapter 7 investigates the finite wordlength effects in our IIR compensation filter including the roundoff noise, coefficient quantization and limit cycle oscillation.

Chapter 8 is the conclusion and future work.

CHAPTER 2 : LITERATURE REVIEW

2.1 TIADC Architecture

The sampling process by an ideal single ADC and an ideal TIADC are briefly reviewed in this section. Throughout this thesis, we shall use Ω to denote the continuous-time frequency variable; $x_a(t)$ and $x[n]$ to denote the continuous-time and discrete-time domain signals respectively; and $X_a(j\Omega)$ and $X(e^{j\Omega T_s})$ to denote the Fourier transforms of $x_a(t)$ and $x[n]$, respectively.

Figure 2.1 An ideal ADC is modelled by a uniform sampler and a quantizer.

An ideal ADC performs two basic functions on an analog signal: 1) samples the analog signal at uniform time intervals and hold it for a specified time and 2) converts the held signal into finite number of discrete levels. As shown in Figure 2.1, an ideal ADC is equivalently modeled by a uniform sampler and a quantizer. Ignoring the quantization effect, the ideal output sequence $\check{x}[n]$ is obtained by sampling the analog input signal $x_a(t)$ uniformly at time instants $t = nT_s, n \in \mathbb{Z}$ where $\mathbb{Z}$ is the space containing all integers, i.e.,

$$\check{x}[n] = x_a(nT_s), \quad n \in \mathbb{Z}, \quad (2.1)$$

where T_s is the sampling period of the ADC. In the frequency domain, the Fourier transforms of $x_a(t)$ and $\check{x}[n]$ are related according to

$$\check{X}(e^{j\Omega T_s}) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} X_a \left(j \left(\Omega - \frac{2\pi k}{T_s} \right) \right), \quad (2.2)$$

where Ω is the continuous-time frequency variable in radians per second (rads/s). Suppose the analog input signal $x_a(t)$ has a frequency spectrum as shown in Figure 2.2(a), then the spectrum of the uniformly sampled sequence $\check{x}[n]$ will be that shown in Figure 2.2(b). Throughout this thesis, it is assumed that 1) the analog input signal $x_a(t)$ is bandlimited to Ω_b rads/s, and 2) The Nyquist criterion is fulfilled so that no aliasing occurs, i.e., $\Omega_b \leq \pi/T_s$. From (2.2), the spectrum of $\check{x}[n]$ is periodic with a period of 2π with respect to ΩT_s . Therefore, if the above two assumptions are valid, the analog signal $x_a(t)$ can be uniquely recovered from $\check{x}[n]$ by considering only one period of $\check{X}(e^{j\Omega T_s})$, i.e.,

$$\check{X}(e^{j\Omega T_s}) = \frac{1}{T_s} X_a(j\Omega), \quad |\Omega| \leq \pi/T_s. \quad (2.3)$$

Figure 2.2 Spectrum of (a) the bandlimited analog signal $x_a(t)$ and (b) the ideally sampled sequence $\check{x}[n]$.

The concept of a TIADC operating at an overall sampling rate of f_s is illustrated in Figure 2.3. The architecture consists of M sub-ADCs working in a round-robin manner. The detailed structure of the sub-ADCs is not of concern as long as all the sub-ADCs are identical. The M sub-ADCs sample and convert the analog input signal at the rate of f_s/M according to the M clock signals $(\phi_0, \phi_1, \dots, \phi_{M-1})$ which are usually generated by a multiphase clock generator. Each clock signal is delayed from the precedent one by T_s seconds, where $T_s = 1/f_s$ is the sampling period of the overall system. Thus, the

M sub-ADCs operate at interleaved time instants as if they were a single ADC operating at the full sampling rate f_s . Digital outputs from all the M sub-ADCs are then combined by the digital multiplexer to form the complete digital stream with data rate of f_s samples per second.

Figure 2.3 Architecture of an M -channel TIADC and the M clock signals driving the sub-ADCs.

Time interleaving is an effective way to implement a high sampling rate ADC without placing too much burden on the device technology. However, TIADCs suffer from errors caused by mismatches among the sub-ADC channels which would not be a problem in a single ADC. The TIADC architecture in Figure 2.3 also includes M sample-and-hold (S/H) circuits in front of the M sub-ADCs. The S/H circuits in an M -channel TIADC must have M -times wider bandwidth than that is required by a single ADC operating at f_s/M ; otherwise, the high frequency components of the input signal will be attenuated.

2.2 Sources of Mismatch Errors and Their Effects

An ideal TIADC consists of M sub-ADC channels with identical characteristics. However, this can never be achieved in practice. Any mismatch between the sub-ADCs introduces unwanted aliasing images and distortion tones in the TIADC output spectrum leading to degraded SFDR and effective resolution of the system. Detailed analysis of the mismatch errors and how they affect the system performance have been discussed in a number of literatures [4-11] and are summarized here.

2.2.1 Sources of Mismatches

a) DC offset mismatch

Each sub-ADC has a constant DC value from the ideal ground level which may be attributed to the DC offset of the op-amp and comparator, charge injection of the switch etc. The DC value can be different for different sub-ADCs leading to a mismatch in the DC offset voltage. Denote the DC offset in the m -th sub-ADC channel as o_m . In the time domain, DC offset is equivalent to adding a constant voltage o_m to the m -th channel output. The offset mismatch is independent of the input signal and manifested at the output stream of the TIADC as a periodic additive noise with the pattern of $\{o_0, o_1, \dots, o_{M-1}, o_0, o_1, \dots, o_{M-1}, \dots\}$.

b) Gain mismatch

Figure 2.4 Transfer characteristics of different ADCs experiencing different gains.

The gain of an ADC is represented by the slope of the best fit line to the ADC transfer characteristic as shown in Figure 2.4. An ideal ADC has unity gain. The gain from the analog input to the digital output may also be different for different sub-ADCs leading to gain mismatch. Gain mismatch among different channels are attributed to different gains of the sample and hold amplifier, inequality of passive components etc. Denote the gain of the m -th sub-ADC as g_m . In the time domain, the effect of gain mismatch is equivalent to multiplying the ideally sampled sequence by a periodic pattern

$\{g_0, g_1, \dots, g_{M-1}, g_0, g_1, \dots, g_{M-1}, \dots\}$ and thus gain mismatch behaves like amplitude modulation. The maximum gain error occurs at the peaks of the input signal.

c) Timing mismatch

Figure 2.5 An example of a two-channel TIADC. (a) Ideal sampling instants (dashed line). (b) Real sampling instants (solid line).

Ideally, the sampling instants of the M sub-ADCs are equally spaced apart by T_s seconds. In practice, Channel- m samples the input signal at $t = kMT_s + mT_s + \Delta t_m, k \in \mathbb{Z}$, where Δt_m is the time skew of Channel- m which is the deterministic deviation between the ideal and real sampling instants. There is also another kind of timing error which is the stochastic deviation from each ideal sampling point named *clock jitter*. *Clock jitter* affects the performance of a single ADC as well and tends to spread out the effect on the whole output spectrum by raising the noise floor. In the context of TIADC, timing mismatch refers to the mismatch in the deterministic time skew Δt_m . Possible sources contributing to the timing mismatch include imperfections of the multiphase clock generation circuit, asymmetric distribution of the clock path, different transistor gate delay, parasitic effects etc. TIADC with timing mismatches produces a periodically nonuniformly sampled sequence. This is illustrated in Figure 2.5 with an example of a two-channel TIADC. The effect of the timing mismatch in time domain is like phase modulation. The maximum of the error introduced by the timing mismatch occurs when the input signal has the largest slew rate.

d) Frequency response mismatch

In addition to the offset, gain and timing mismatches introduced above, a TIADC may also experience frequency response mismatch. The input behavior of each sub-ADC channel can be characterized by a linear system, of which the frequency response may differ from that of the other sub-ADCs'. Denote the channel transfer function of the m -th channel as $G_m(j\Omega)$. A special case of the frequency response which is widely studied is the bandwidth mismatch. The S/H circuit in sample mode can be approximated by a first order RC circuit constituted of the on-resistance of the MOS sampling switch and the hold capacitance. Due to the fabrication process imperfections, components mismatch is inevitable resulting in mismatched bandwidths. Denote the bandwidth of the S/H in the m -th channel as $\Omega_{c,m}$ rads/s, the channel transfer function of the m -th channel can be approximated by

$$G_m(j\Omega) = \frac{1}{1 + j \frac{\Omega}{\Omega_{c,m}}} . \quad (2.4)$$

It has been derived in [4] that for an input signal with frequency Ω_{in} , bandwidth mismatch introduces additional gain and timing mismatches given by

$$g_{c,m} = 1 / \sqrt{1 + (\Omega_{in} / \Omega_{c,m})^2} , \quad (2.5)$$

$$\Delta t_{c,m} = - \arctan(\Omega_{in} / \Omega_{c,m}) / \Omega_{in} . \quad (2.6)$$

$g_{c,m}$ and $\Delta t_{c,m}$ in (2.5) and (2.6) depend on the input frequency Ω_{in} and are different from the gain and time skew discussed before. Bandwidth mismatch causes more severe errors at higher input frequency. The static time skew Δt_m may be regarded as a special case of channel frequency response. An amount of Δt_m deviation from the ideal sampling edge is equivalent to sampling a time shifted version of the input signal $x_a(t + \Delta t_m)$ at the ideal sampling instants, i.e., the channel transfer function is $G_m(j\Omega) = e^{j\Omega\Delta t_m}$.

2.2.2 Effects of the Mismatches in the Frequency Domain

In order to get a clear picture of how the combined mismatches interact and affect the frequency spectrum of the TIADC output, a simple mathematical model of the m -th sub-ADC channel as shown in Figure 2.6 is assumed. The effects of gain, time skew and frequency response mismatch in Channel- m are modeled by the linear system $G_m(j\Omega)$. Offset mismatch cannot be modeled by $G_m(j\Omega)$ and is represented by the addition of the DC voltage o_m . The input signal first passes through the linear system $G_m(j\Omega)$, then added with a DC offset o_m and finally sampled with a sampling interval of MT_s . The sampling instants of Channel- m are delayed by mT_s from that of Channel-0.

Figure 2.6 An equivalent model of the m -th sub-ADC channel.

Neglecting the quantization noise, the output sequence of the m -th sub-ADC channel is written as

$$y_m(t) = \sum_{k=-\infty}^{\infty} (x_a(t) * g_m(t) + o_m) \cdot \delta(t - kMT_s - mT_s), \quad (2.7)$$

where $g_m(t)$ is the continuous-time impulse response of $G_m(j\Omega)$, $\delta(t)$ is the Dirac delta function and “ $*$ ” denotes the convolution operation. The TIADC output is obtained by summing the output sequence of all the M sub-ADCs. Because the input signal is bandlimited and the Nyquist criterion is fulfilled and the TIADC output spectrum is periodic with period Ω_s , where $\Omega_s = 2\pi f_s$, it suffices to consider only one period of the output spectrum, from 0 to Ω_s , then

$$Y(j\Omega) = \frac{1}{T_s} \sum_{k=0}^{M-1} \left[A_k \left(j \left(\Omega - \frac{k\Omega_s}{M} \right) \right) X_a \left(j \left(\Omega - \frac{k\Omega_s}{M} \right) \right) + 2\pi B_k \delta \left(\Omega - \frac{k\Omega_s}{M} \right) \right], \quad (2.8)$$

where

$$A_k(j\Omega) = \frac{1}{M} \sum_{m=0}^{M-1} G_m(j\Omega) e^{-j\frac{2\pi km}{M}}, \quad (2.9)$$

$$B_k = \frac{1}{M} \sum_{m=0}^{M-1} o_m e^{-j\frac{2\pi km}{M}}. \quad (2.10)$$

Several observations can be made from (2.8-2.10).

- 1) The gain, timing and frequency response mismatches interact with each other and generate spectral components which are copies of the input signal weighted by $A_k(j\Omega)/T_s$ shifted by integer multiples of the channel sampling rate, i.e., $k\Omega_s/M$, $k = 1, 2, \dots, M-1$ whereas offset mismatch creates spectral lines with magnitudes of $2\pi B_k/T_s$ located at $k\Omega_s/M$, $k = 0, 1, \dots, M-1$.
- 2) The desired signal term is $A_0(j\Omega)X_a(j\Omega)$ where $A_0(j\Omega)$ is a frequency-dependent gain term introduced to the original signal.
- 3) The *distortion weighting coefficient* sets $A_k(j\Omega)$ and B_k are discrete Fourier transforms (DFTs) of their corresponding channel mismatch parameters $G_m(j\Omega)$ and o_m , respectively. $A_k(j\Omega)$ and B_k are periodic with period M with respect to k .
- 4) Using the DFT property, the spectral lines introduced by DC offset mismatch have the symmetric property that $[B_k]^* = B_{-k} = B_{M-k}$ where $[]^*$ denotes the complex conjugate of $[]$. The distortion coefficients A_k introduced by gain mismatch have the same property that $[A_k]^* = A_{-k} = A_{M-k}$. For timing and frequency response mismatches, it is easy to verify that $A_k(-j\Omega) = [A_{M-k}(j\Omega)]^*$. To get a straightforward view, the output spectrum of a four-channel TIADC assuming a sinusoidal input signal $x_a(t) = a \cdot \sin(\Omega_{in}t)$ is shown in Figure 2.7. The solid line represents the desired signal components and the dashed line represents the spectral images and distortion tones caused by the mismatches. The amplitude of the input signal is omitted for simplicity. Figure 2.7(a) plots the spectrum when only offset and gain mismatches are present. The distortion coefficients are independent of the input frequency. The images introduced by gain mismatch have the same magnitude for each pair of negative and positive frequency components, i.e., the distortion pair is symmetric about

$k\Omega_s/M$. The output spectrum in the presence of timing and frequency response mismatches is shown in Figure 2.7(b). The distortion coefficients are frequency dependent and the distortion pairs are in general not symmetric about $k\Omega_s/M$.

If all the channels have identical frequency characteristic $\bar{G}(j\Omega)$ and identical offset voltage $\bar{o}$, the coefficient sets $A_k(j\Omega)$ and B_k are then reduced to

$$A_k(j\Omega) = \frac{1}{M} \bar{G}(j\Omega) \sum_{m=0}^{M-1} e^{-j\frac{2\pi km}{M}} = \begin{cases} \bar{G}(j\Omega), & k = 0, M, 2M, \dots \\ 0, & \text{otherwise} \end{cases}, \quad (2.11)$$

$$B_k = \frac{1}{M} \bar{o} \sum_{m=0}^{M-1} e^{-j\frac{2\pi km}{M}} = \begin{cases} \bar{o}, & k = 0, M, 2M, \dots \\ 0, & \text{otherwise} \end{cases}. \quad (2.12)$$

In this case, the aliasing components no longer exist and the output spectrum is the same as that of a single ADC with frequency characteristic $\bar{G}(j\Omega)$ and DC offset $\bar{o}$.

Figure 2.7 Output spectrum of a four-channel TIADC over $[0, \Omega_s]$ assuming a sinusoidal input with frequency Ω_{in} . (a) Gain and offset mismatches. (b) Timing and frequency response mismatches.

The distortion components introduced by the mismatches degrade the performance of the TIADC such

as the SINAD and SFDR. The SINAD due to individual mismatch errors can be explicitly derived using either the given mismatch parameters or parameter distributions [5-11]. Assume the same input signal $x_a(t) = a \cdot \sin(\Omega_{in}t)$ and ignore the quantization noise. If the offset values are explicitly known, the SINAD due to offset mismatch only can be calculated as

$$\text{SINAD} = 10\log_{10} \left(\frac{a^2/2}{1/M \cdot \sum_{m=0}^{M-1} o_m^2} \right). \quad (2.13)$$

If the channel offsets $o_m, m = 0, 1, \dots, M-1$ are M independent, identically distributed (*i.i.d.*) random variables with zero mean and standard deviation σ_o , the expected SINAD is then ([5-7])

$$\text{SINAD} = 20\log_{10} \left(\frac{a}{\sqrt{2}\sigma_o} \right). \quad (2.14)$$

The SINAD due to offset mismatch is independent of the number of interleaving channels and can be improved by 6 dB for every doubling of the input amplitude or halving of the offset mismatch. When the gain mismatch is the sole mismatch source and the values are available, the SINAD is

$$\text{SINAD} = 10\log_{10} \left(\frac{(\sum_{m=0}^{M-1} g_m)^2}{M \sum_{m=0}^{M-1} g_m^2 - (\sum_{m=0}^{M-1} g_m)^2} \right). \quad (2.15)$$

Suppose the channel gain $g_m, m = 0, 1, \dots, M-1$ are M *i.i.d.* random variables with mean μ_g and standard deviation σ_g , the expected SINAD is ([5-7])

$$\text{SINAD} \approx 20\log_{10} \left(\frac{\mu_g}{\sigma_g} \right) - 10\log_{10} \left(1 - \frac{1}{M} \right). \quad (2.16)$$

The SINAD due to gain mismatch is affected by the number of interleaving channels, but the effect is very limited. When M is increased from two to infinity, the SINAD degrades by 3 dB only. With known timing mismatch, the SINAD due to timing mismatch can be calculated according to

$$\text{SINAD} = 10\log_{10} \left(\frac{|\sum_{m=0}^{M-1} e^{j\Omega_{in}\Delta t_m}|^2}{M^2 - |\sum_{m=0}^{M-1} e^{j\Omega_{in}\Delta t_m}|^2} \right). \quad (2.17)$$

Suppose the time skew $\Delta t_m, m = 0, 1, \dots, M - 1$ are M *i.i.d.* Gaussian random variables with zero mean and standard deviation σ_τ , the expected SINAD can be approximated by ([5-9])

$$\text{SINAD} \approx 20\log_{10}\left(\frac{1}{\Omega_{in}\sigma_\tau}\right) - 10\log_{10}\left(1 - \frac{1}{M}\right). \quad (2.18)$$

The expected SINAD does not depend on the actual distribution of the offset and gain mismatch, only their respective means and standard deviations, whereas the distribution of the timing mismatch is necessary to calculate the expected SINAD. The distortion caused by timing mismatch becomes more severe as the input frequency increases. For every decade increase in the input frequency or the time skew standard deviation, the SINAD drops by 20 dB. SINAD analysis when only bandwidth mismatch is provided in [11] by using the first order model given in (2.4). Define a new variable ϵ_m as $\epsilon_m = \Omega_{ref}/\Omega_{c,m} - 1$ where Ω_{ref} is the reference bandwidth. By assuming $\epsilon_m, m = 0, 1, \dots, M - 1$ are M *i.i.d.* Gaussian random variables with zero mean and standard deviation σ_ϵ , the expected SINAD due to bandwidth mismatch is derived as

$$\text{SINAD} \approx 10\log_{10}\left(1 + \frac{\Omega_{in}^2}{\Omega_{ref}^2}\right) - 20\log_{10}\left(\frac{\sigma_\epsilon\Omega_{in}}{\Omega_{ref}}\right) - 10\log_{10}\left(1 - \frac{1}{M}\right). \quad (2.19)$$

Similar to the case of timing mismatch, SINAD due to bandwidth mismatch is inversely proportional to the square of the input frequency.

2.3 Mismatch Estimation and Calibration

The mismatches caused by component inequalities among different sub-ADCs can be reduced by laser trimming technology. However fabrication process can be rather costly and time consuming; component matching becomes more difficult to achieve with ever shrinking IC technologies and increasing clock speed. Further, the mismatch parameters may drift due to temperature variation or component aging. Trimming cannot track variations over time. Post fabrication methods which relax the stringent requirements on the IC technologies have been proposed by a number of researchers to

attenuate the spurious spectral components. It may be done by either adding extra analog circuitry to the ADC's front end or post-processing the digital output samples. The variations of the mismatch parameters due to environmental change are usually very slow and thus can be tracked and corrected by adaptive algorithms. In this section, the correction and identification methods for each of the four types of mismatches, namely the offset, gain, timing and frequency response mismatches are reviewed.

2.3.1 Offset Mismatch

The DC offset may be estimated by calculating the mean value of a batch of sub-ADC output samples [12] or measured at the digital output with zero input signal [13-14]. Either analog or digital method may be employed for offset correction. Analog correction can be done by for example, connecting an auxiliary capacitor storing the offset voltage to the input of the respective sub-ADC. Alternatively, digital correction method may be employed where the estimated offset value is subtracted from the output samples of the corresponding sub-ADC. Analog method does not reduce the signal range whereas digital method is more robust. A digital averaging method for a two-channel TIADC was employed in [15] where the output of each channel was delayed by one sample and added to the other channel. The drawback is that the digital averaging has a low pass bandwidth of 1/2 of the Nyquist frequency. The mismatch estimation method presented in [16] requires an extra reference ADC. The gain and offset of the sub-ADC are calculated from the outputs after applying identical clock and analog input signal to the sub-ADC and the reference ADC.

Figure 2.8 The LMS loop for analog background tracking and correction of the offset and gain mismatches in [17].

Background adaptive methods which do not interrupt the normal operation of the TIADC and are able to track slowly varying mismatches have been presented in [17-23]. The analog background calibration method demonstrated in [17-18] requires $M + 1$ conversion ADCs and a reference ADC for a M -channel TIADC. At any time, M sub-ADCs do the conversion while the extra sub-ADC, denoted as ADC- i and the reference ADC are fed identical calibration signals. A least mean squares (LMS) loop as shown in Figure 2.8 is then employed to track and correct the offset in ADC- i over time. In each update, the output difference between the reference ADC and ADC- i is added to ADC- i input and in steady-state the offset of ADC- i equals to that of the references ADC's. When the calibration is done, another sub-ADC in the M -channel array is selected for calibration and ADC- i is switched back to perform normal conversion. The disadvantages of this method are the need of two extra ADCs and pattern noise due to periodically swapped in/out of sub-ADCs. A digital background calibration method for a two-channel TIADC which avoids these problems was introduced in [18-19]. The same LMS algorithm as that shown in Figure 2.8 was employed except that the error is between the output of the two sub-ADCs and the accumulator output is added to the output digital code of the corresponding sub-ADC. One drawback associated with this digital offset estimation method is that the input signal cannot have frequency components located at $kf_s/2, k \in \mathbb{Z}$ because signal at these frequencies after being sampled by different sub-ADCs appear as different DC voltages and cannot be distinguished from the offset. The problem associated with the frequency component at $kf_s/M, k \in \mathbb{Z}$ was overcome in [20-22] where the input signal is multiplied by an uncorrelated pseudo random sequence consisting of only two values, +1 and -1. We denote this sequence as $PRN[n]$ in this chapter. The resulting analog signal becomes a white noise with zero mean and thus has little information at DC. The offset value is then obtained by averaging [20, 22] or low pass filtering [21] and subtracted from the corresponding sub-ADC output. The signal is recovered by multiplying the offset removed sequence with the same pseudo random sequence. The online equalization technique presented in [23] uses a reference ADC and an LMS loop to estimate the offset and gain mismatch parameters and a linear equalizer to correct for the mismatches. The techniques presented in [24-25] have exploited existing system hardware to correct for offset mismatch of the TIADC used in a communication system. In [24], the signal presented at the ADC input has a wide-band spectrum and

no DC content. The output sample averaged over time corresponds to the offset value and the LMS algorithm was employed for correction.

A fast Fourier transform (FFT) based offset correction method was presented in [26]. The method calculates the FFT of a batch of output samples and estimates the offset values from the mismatch distortion power. The offsets are then removed from the output samples and the FFT of the new data is evaluated again. This process continues until the mismatch distortion power falls below some threshold. A blind estimation method has been demonstrated in [27] where the input signal is assumed to be quasi-stationary and the offset values are estimated by iteratively minimizing a cost function parameterized on the DC offsets. The approaches in [26] and [27] require a substantial amount of computation due to the repetitive FFT evaluation and gradient searching.

2.3.2 Gain Mismatch

Gain matching may be improved offline by using components with better matching, careful layout design or hardware sharing [12-14]. Alternatively, gain error can be corrected in the background using an LMS adaptive loop shown in Figure 2.8 [17-18]. The negative feedback loop forces the gain of ADC- i equal to that of the reference ADC in the steady state. The authors of [18-19] have presented a digital background calibration technique which employs the LMS loop shown in Figure 2.9. An uncorrelated pseudo random sequence $PRN[n]$, after a 1-bit digital-to-analog converter (DAC) is added to the input of ADC- i . The same random sequence is subtracted from the gain corrected digital output to restore the original signal. The variable gain in ADC- i is updated adaptively according to

$$g_i[n + 1] = g_i[n] - \mu \cdot PRN[n] \cdot e_i[n], \quad (2.20)$$

where μ is the updating step size. The error signal $e_i[n]$ in (2.20) consists of two part, one is the input signal multiplied by the gain of ADC- i , g_{ADC} , and the variable gain, $g_i[n]$; the other is the $PRN[n](g_{DAC}g_{ADC}g_i[n] - 1)$. In the long run, the accumulator output contributed by the first part zeros out and the variable gain makes the overall gain in ADC- i equal to $1/g_{DAC}$.

The authors of [22] demonstrated that multiplying the output of a two-channel TIADC with a chopping signal which alternates between $+1$ and -1 at the rate of f_s causes the signal and spur to switch position in the spectrum. The product of the chopped output and the TIADC output accumulated over time is a DC term proportional to the relative gain between the two channels. A direct extension of the chopper-based method to a four-channel TIADC was presented in [28]. A related technique which estimates and compensates for the gain mismatch by applying the Hadamard transform to the TIADC output was presented in [29]. In the case of a two-channel TIADC, the gain estimation methods of [22] and [29] are equivalent as the Hadamard transform of order two has the same function as a chopper. A problem associated with the techniques of [22, 28-29] is that when the input signal has components at $kf_s/2M, k \in \mathbb{Z}$, the product of the TIADC output and the chopped output is a DC term even when there is no gain mismatch.

Figure 2.9 The LMS loop for digital background tracking and correction of the gain mismatch in [19].

The method presented in [24] caters for the TIADCs used in digital communication systems. The correlation between the channel output and the slicer error in that channel is used to update the variable gain connected to the output of the sub-ADC. In [30], the input signal is multiplied by a pseudo random sequence alternating between 0 and 1 and the digital output is correlated with the product of the same random sequence and the input in the background to extract the gain. An FFT-based method is found in [26] and a blind adaptive method is found in [27]. Randomizing the selection of the sub-ADCs for each sample transforms the offset and gain mismatch induced spurs into noise like distortions and has been employed to improve the SFDR performance in [31-32].

2.3.3 Timing Mismatch

Timing mismatch tends to cause more severe error as the input frequency increases. Timing mismatch introduced by the multi-phase clock generator may be controlled by careful circuit design prior to fabrication [12]. However, the improvement of the accuracy is very limited and the time skew variations due to environmental change cannot be tracked. The techniques for timing mismatch correction so far may be classified into two broad categories, one is analog or mixed signal correction [33-49] and the other is digital interpolation [52-77].

Figure 2.10 A two-rank S/H system to avoid timing mismatch caused by nonideal sampling instants.

Mismatch in the time skews can be avoided by using a single front-rank S/H in front of all the interleaving channels [13-14, 17-19, and 33-35]. In a two-rank S/H system, the single full speed S/H operating at f_s acquires samples on the analog input and the second-rank S/H reacquires and digitizes the held sample. Second-rank S/H circuits go from track to hold during the hold time of the first-rank S/H when the voltage at the inputs of the second-rank S/Hs is kept constant. Therefore, the timing mismatch caused by the nonideal sampling instants has no effect. Figure 2.10 shows a two-rank S/H structure. In Figure 2.10, ϕ_s is the clock signal driving the front-rank S/H and $\phi_m, m = 0, 1, \dots, M - 1$ are the clock signals driving the second-rank S/Hs. Despite the clear advantage offered by the two-rank S/H scheme, the requirement that the front-rank S/H must operate at the full sampling rate

imposes stringent constraints on the settling time of the S/H, making the circuit design a great challenge in practice. Therefore the number of channels that may be interleaved is limited.

Mixed-signal approaches which detect the timing information from the output samples using digital processing and adjust the analog circuitry to eliminate the timing mismatch can be found in a number of works [9, 24, 36-49]. In the foreground techniques, a special phase where a known signal is injected for time skew identification is needed at start up or when the time skew drifts and then a variable delay circuit is used to adjust the channel sampling clocks [9, 36-38]. On the other hand, background methods estimate and correct for the timing mismatch at the same time by employing LMS-based algorithms [24, 39-46] or by using correlation functions [47]. The estimated timing information is used to adaptively adjust the digitally controlled delays connected to the clock buffers until the mismatch is minimized. In [49], the effects of timing and bandwidth mismatches were alleviated simultaneously by tuning the transfer characteristics of the S/Hs such that the phase mismatch introduced by the S/Hs counteracts the timing mismatch.

Using analog circuits for mismatch correction has several drawbacks such as low accuracy and difficulty in porting to other process technologies. Another option to calibrate the timing mismatch without adjusting the analog circuitries is by processing the TIADC raw output in the digital domain to interpolate the sample values that would have been captured at the ideal sampling instants [52-77]. Digital processing offers guaranteed accuracy, great reliability and flexibility and becomes faster and less expensive as the process technology downscales.

A well known theorem for the reconstruction of a band-limited continuous-time signal $x_a(t)$ from its nonuniform samples by Yao and Thomas [50] is that, if a band-limited continuous-time signal, $x_a(t)$, is sampled at an irregular grid $\{t_k, k \in \mathbb{Z}\}$ with an average rate exceeds the Nyquist rate, then $x_a(t)$ is uniquely determined by its nonuniform samples $x_a(t_k)$ according to

$$x_a(t) = \sum_{k=-\infty}^{\infty} x_a(t_k) \frac{G(t)}{G'(t_k)(t - t_k)}, \quad (2.21)$$

where

$$G(t) = (t - t_0) \prod_{k=-\infty, k \neq 0}^{\infty} \left(1 - \frac{t}{t_k}\right). \quad (2.22)$$

In (2.21), $G'(t_k)$ is the derivative of $G(t)$ evaluated at $t = t_k$, and t_0 can be chosen to be zero. In the case of a TIADC, $t_k = kT_s + \Delta t_{\langle k \rangle M}$ where $\langle k \rangle M$ denotes the k modulo M operation. By incorporating (2.21) and the periodical sampling pattern of the TIADC, Eldar and Oppenheim [52] first derived a continuous-time synthesis filter bank and then an equivalent discrete-time synthesis filter bank to interpolate the uniform samples. The ideal filter derived has a piecewise constant frequency response and achieves perfect reconstruction. However, the practical case of designing casual filters of finite order was not addressed.

Figure 2.11 (a) An all digital multirate filter bank structure for the reconstruction of uniform samples from the nonuniformly sampled TIADC output. (b) An equivalent hybrid model with continuous-time analysis bank and discrete-time synthesis bank.

The techniques presented in [52-58] for reconstructing the uniform samples have the form of a multirate filter bank structure [51] as shown in Figure 2.11(a). The discrete time sequence $\tilde{x}[n]$ represents the samples of the analog input signal $x_a(t)$ taken uniformly at rate f_s . The analysis filter bank consists of the discrete-time filters $G_m(e^{j\Omega T_s}) = e^{j\Delta r_m \Omega T_s}$, $m = 0, 1, \dots, M-1$ which represent the time skews in the sub-ADC channels; $G_m(e^{j\Omega T_s})$ is the discrete-time equivalent of $G_m(j\Omega)$ and Δr_m is the time skew normalized with respect to the overall sampling period, i.e., $\Delta r_m = \Delta t_m / T_s$. The term $e^{jm\Omega T_s}$ models the desired uniform delay among the channel sampling clocks. The output of the M sub-ADCs are the M sequences after the M -fold down-sampler. The reconstruction is performed by the right half of the multirate filter bank, including the M -fold up-sampler, the delay term $e^{-jm\Omega T_s}$ and the synthesis filter bank $H_m(e^{j\Omega T_s})$ where $m = 0, 1, \dots, M-1$. The underlying principle of the multirate filter bank approach can be summarized as follows: the output of each sub-ADC channel is an under sampled and thus aliased version of the analog input signal $x_a(t)$; when the sub-sequences are up-sampled by M and summed to form the final output, the aliasing terms persist due to mismatch error; synthesis filter banks are designed to shape the spectral characteristics of the output of each sub-ADC in such a way that the aliasing components are minimized whereas the signal term is retained as much as possible when the filter outputs are summarized together. The analysis filter bank in Figure 2.11(a) does not exist in actual implementation. An equivalent hybrid model with continuous-time analysis bank modeling the non-ideal sampling and discrete-time synthesis bank modeling the reconstruction is shown in Figure 2.11(b). The time skew in each channel is represented by the continuous-time transfer function $G_m(j\Omega) = e^{j\Delta t_m \Omega}$. Ignoring the quantization effect, and using the same derivation as in Section 2.2.2, the Fourier transform of the final output sequence $y[n]$ in one period can be written as

$$Y(e^{j\Omega T_s}) = \frac{1}{T_s} \sum_{k=0}^{M-1} V_k(j\Omega) X_a \left(j \left(\Omega - \frac{2\pi k}{MT_s} \right) \right), \quad (2.23)$$

where

$$V_k(j\Omega) = \frac{1}{M} \sum_{m=0}^{M-1} e^{-j\frac{2\pi km}{M}} e^{j\left(\Omega - \frac{2\pi k}{MT_s}\right)\Delta t_m} H_m(e^{j\Omega T_s}). \quad (2.24)$$

$V_0(j\Omega)$ represents the frequency dependent gain imposed on the input signal and $V_k(j\Omega)$ where $k = 1, 2, \dots, M - 1$ represent the gain of the aliasing components. We call them *signal function* and *aliasing functions* respectively. The condition that $V_0(j\Omega) = ce^{-jd\Omega T_s}$ and $V_{k \neq 0}(j\Omega) = 0$ over the frequency range of $|\Omega| \leq \pi/T_s$ is called *perfect reconstruction* (PR) [51] where c is a constant and d represents an integer delay. In the PR condition, the output is simply a delayed and scaled version of the ideally sampled sequence, i.e., $y[n] = c\tilde{x}[n - d]$. The synthesis filters $H_m(e^{j\Omega T_s})$ in [53-55] are in the form of $H_m(e^{j\Omega T_s}) = a_m ce^{-j\Omega T_s(\Delta r_m + d)}$. The gain constant a_m is chosen such that the *signal function* $V_0(j\Omega) = ce^{-jd\Omega T_s}$ and *aliasing function* $V_{k \neq 0}(j\Omega) = 0$ and the digital fractional delay filters are approximated by FIR filters. Oversampling is required by the proposed method; $x_a(t)$ has to be bandlimited by $\Omega_b \leq \pi(M + 1)/2MT_s$ for odd M and $\Omega_b \leq \pi/2T_s$ for even M . The authors of [56] have approximated the linear phase terms $e^{j(m + \Delta r_m)\Omega T_s}$, $m = 0, 1, \dots, M - 1$ by an FIR filter bank in polyphase matrix form and derived the FIR synthesis filter bank using the PR condition in polyphase form [51]. Oversampling is required to alleviate the effects of inaccuracies in approximating the analysis filter bank. A similar approximation can be found in [57] and the design of the synthesis filter bank was viewed as a communication equalization problem. The authors of [58] assumed a stationary known input spectrum and designed compensation system in the form of a Wiener filter using a statistically optimal LMS criterion. If infinite order is allowed for the filters in [56] and [58] and if the input is restricted to be a bandlimited signal in [58], the methods of [56] and [58] achieves PR and are equivalent to the ideal filters derived in [52]. The compensation structures in [22, 28] took the form of multirate filter bank as well where the FIR synthesis filters were designed by making $V_0(j\Omega) = 1$ and $V_{k \neq 0}(j\Omega) = 0$. Either closed-form ([22]) or approximated solutions ([28]) were obtained.

In [59-63], signal reconstruction was performed by employing a discrete-time PTV FIR filter according to

$$y[n] = \sum_{p=-N}^N x[n - p]h_n[p], \quad (2.25)$$

where $h_n[p]$ represents the impulse response of the PTV FIR filter and $h_n[p] = h_{n+M}[p]$. Utilizing the periodicity, the PTV filter reconstruction system was equivalently represented by an all-digital

multirate filter bank structure where the analysis filter bank is in the form of $G_m(e^{j\Omega T_s}) = \sum_{p=-N}^N h_m[p]e^{-j\Omega T_s(p-\Delta r_{m-p})}$ and all the synthesis filters are equal to one. Then the filter coefficients $h_m[p], m = 0, 1, \dots, M-1$ were optimally designed in a minimax sense such that the *signal function* $V_0(j\Omega)$ and the *aliasing function* $V_{k \neq 0}(j\Omega)$ approximate one and zero respectively. The resulting filter order is lower than that in [56] for the same level of reconstruction accuracy.

An improved structure which avoids online filter design when the timing error in a two-channel TIADC varies was presented in [60]. The design utilizes the Farrow structure [78] where the coefficients $h_1[p]$ are expressed in polynomials of Δr_1 where Δr_1 is the relative timing mismatch in Channel-1 with reference to Channel-0. The resulting implementation is a linear combination of a few fixed sub-filters weighted by different powers of Δr_1 . Straightforward extension to the M -channel case leads to the use of polynomials in multiple variables, $\Delta r_m, m = 1, 2, \dots, M-1$ and thus the design complexity would increase rapidly as M grows [61]. The methods presented in [62] and [63] reduced the online design complexity by using a two-rate based technique where the design is splitted into two parts: one is offline design of a linear phase half-band filter catering for all combinations of the extreme cases of the time skews; and the other is online design of a PTV FIR filter with smaller complexity. The offline design part requires solving a number of nonlinear optimization problems.

The methods reported in [29, 64-70] have assumed a common model where the TIADC output is decomposed into an error signal that is caused by the timing mismatch and a desired signal which is free from the timing mismatch. Compensation is achieved by reproducing the error signal and subtracting it from the TIADC output to obtain the desired signal. In [29], Hadamard transform was employed to generate pseudo aliasing images. The resulting implementation structure comprises only a fixed FIR filter and a few general multipliers for arbitrary number of interleaving channels. In the two-channel case presented in [64], the time skew $e^{j\Delta r_1 \Omega T_s}$ is expressed as a polynomial of Δr_1 using the first three terms of the Taylor series expansion and the error spectrum is reproduced by applying a differentiator to the TIADC output followed by frequency shifting of π . In [65], the nonuniformly sampled signal $x(mT_s + \Delta r_m T_s)$ is expressed as a polynomial of Δr_m . The resulting implementation structures of [64] and [65] consists of FIR filters approximating the ideal differentiator and time-

varying multipliers corresponding to different orders of Δr_m . More stages which corrects for higher order terms may be used for better performance and the number of FIR filters comprised in each stage scales linearly with the number of correction stages. The work in [66-68] employs a differentiator filter, a Hilbert transform filter and a scaling factor to compute the derivative of the input in any Nyquist band and are extensions of [29]. The aliasing signal in [66-67] is generated by passing the derivative through Hadamard transform whereas that in [68] is generated using an equivalent polyphase structure of the TIADC.

A TIADC with timing mismatch is a linear time-varying system and the input-output relation may be represented by a system of linear equations

$$\mathbf{x} = \mathbf{A}\ddot{\mathbf{x}}, \quad (2.26)$$

where $\mathbf{x} = [x[0], \dots, x[N-1]]^T$ and $\ddot{\mathbf{x}} = [\ddot{x}[0], \dots, \ddot{x}[N-1]]^T$ are the output vector and input vector, respectively, of the TIADC and the elements in matrix $\mathbf{A}$ are the coefficients of the time-varying FIR filters approximating the timing mismatch $e^{j\Delta r_m \Omega T_s}$. Utilizing (2.26), the authors of [71] have developed a constrained least squares formulation by incorporating the bandlimited constraint of the input signal. Oversampling is mandatory and the inversion of large dimension matrix leads to considerable amount of computation. Alternatively, an iterative framework where $\ddot{\mathbf{x}}$ is reconstructed in a sample-by-sample manner using Richardson iteration may be employed [72].

The uniform samples may be estimated by using a weighted combination of the neighboring nonuniform samples [73-74]. In [73], the sample time error was corrected by employing a linear casual interpolation method based on an iterative algorithm called “*Neville’s method*”. The compensation structure in [74] is an FIR approximation of the Lagrange interpolation polynomial where the weight of each nonuniform sample, or the filter coefficient is computed using the Lagrange interpolation formula. Lagrange interpolation for timing mismatch correction may also be found in [75] where an exact solution is derived and hence the method in [75] is more accurate than that in [74] when the timing error is large. The author of [76] and [77] has demonstrated a frequency domain approach which recovers spectral information at discrete frequency points of a uniformly sampled

signal based on points in the nonuniformly sampled spectrum. The method requires the processing of a batch of time domain samples to obtain the information of the nonuniformly sampled spectrum, thus preventing its use in real-time applications.

It has been demonstrated in [79-83] that the impact of the timing mismatch can be alleviated by spectral shaping methods which do not rely on accurate estimation of the time skews. One way to shape the spectrum of the TIADC is through randomizing the sampling sequence of the sub-ADCs [79-81]. Extra sub-ADCs need to be introduced to the interleaved structure in order to enable randomization. As the number of additional sub-ADCs increases, the degree of randomness is increased and the spurious images are transformed into a more noise-like distortion. The SFDR is improved but the SINAD remains unchanged since the total noise power remains the same. Doubling the oversampling ratio and using a low pass filter increase the SINAD by 3 dB. A simple sorting algorithm was employed in [82] to control the sampling sequence of the sub-ADCs such that the spurious images residing at higher frequencies near $\Omega_s/2$ have greater power. The out-of-band mismatch spurs were then suppressed by a low pass filter leading to improved SINAD. A method which combines both randomization and the sorting algorithm was presented in [83].

The compensation systems have to be redesigned to cope with the drifts in timing mismatches. It is desirable that the redesign is avoided or at least can be done easily online so that the operation of the TIADC is not halted. The filter coefficients may be updated in the background whenever a new estimate of the time skew is available by using a look up table (LUT) [22] or an adaptive filter bank [28]. Farrow-based structures where the filter coefficients are expressed in polynomials of the time skew values were employed in [53, 60-61, 72]. The resulting structure consists of a number of fixed sub-filters weighted by different powers of the time skews. For the time-varying filter structure in [60-61], Farrow structure is limited to small number of interleaving channels and small time skews due to the high design complexity as M increases. The solutions in [64] and [65] are based on Taylor expansion with respect to the time skews. The compensating structures consist of fixed FIR filters approximating the differentiator and time-varying multipliers corresponding to different orders of the time skew values.

The time skews must be known beforehand when designing the reconstruction filters using the aforementioned techniques. In the foreground detection schemes presented in [9] and [84], a known sinusoidal signal was applied at the input and the time skews were extracted from the distortion tones using inverse DFT (IDFT). Foreground estimation methods usually offer superior accuracy due to dedicated characterization and calibration hardware. Therefore, they are suitable for applications where high resolution is required and system stoppage is tolerable. In the background technique presented in [73], a ramp signal with known slope and period T_s was mixed with the input signal to measure the time skews. The drawbacks associated with this method include accurate ramp signal generation and reduced system dynamic range. The scheme of [85] utilizes the ramp signal for time skew estimation as well. An extra reference ADC and complex channel switching are required. The authors of [24, 39-41] have demonstrated a background time skew detection scheme catered for the digital communication system. The time skew in ADC- i is estimated using a correlation function between the samples of ADC- $(i-1)$ after the slicer and the samples of ADC- i before the slicer. The difference in the correlation function between two adjacent channels is proportional to the time skew mismatch. The principle of the time skew detection scheme in [42-44] is based on the observation that the probability of the zero-crossing between two adjacent samples is proportional to the time skew between the two samples. The zero-crossing detection is sensitive to comparator offset and the clock chopper employed in the detection scheme may introduce additional time skew. The background method in [45] requires an extra single-bit calibration ADC. An LMS-based algorithm is used to adjust the clock delay line so that the cross-correlation between the output of the sub-ADC and the calibration ADC is maximized at convergence.

Blind methods do not rely on accurate knowledge of the input signal and perform timing mismatch estimation and correction in the background while the TIADC is running continuously [22, 27-29, 86-103]. For the blind methods proposed so far, these are usually achieved at the cost of higher computational complexity, being limited to smaller number of channels, and certain restrictions on the input signal.

An expectation-maximize algorithm was employed in [86] to iteratively produce the maximum

likelihood estimates of gain and timing parameters. The input signal is restricted to stationary bandlimited zero-mean Gaussian process. In [87-88], time skews were estimated by minimizing the out-of-band energy of the reconstructed signal using a least squares formulation. Several iterations may be needed to improve the estimation accuracy for large time skews. The techniques can be extended to large number of interleaving channels and implemented using adaptive filtering schemes. The methods in [86-88] requires considerable amount of computation.

Small amount of oversampling ensures that there exists a frequency band around the zero frequency where no aliasing component appears for a two-channel system [89] and only three aliasing components appear for a four-channel system [90]. These spectral properties are then exploited to derive a synthesis filter bank parameterized on the time skews using adaptive null-steering algorithm. The techniques in [89-90] have reduced computational complexity but are restricted to system with a small number of sub-ADCs.

Blind identification methods presented in [91-94] have exploited the spectral information in the signal free band centered at $\Omega_s/2$. By expressing the spectral content in the signal free region in terms of the input signal and the *distortion weighting function* $A_k(j\Omega)$, $A_k(j\Omega)$ is obtained by using a least squares formulation and then the time skews are identified using IDFT [91]. The time skew estimate in a two-channel TIADC is updated by using an LMS algorithm so that the difference between the actual error signal in the signal free region obtained by high pass filtering and the reproduced error signal parameterized on the estimate is minimized [92]. An extension to the identification of both gain and timing mismatch in an arbitrary number of channels was provided in [93]. The method presented in [94] employed a filtered-error LMS algorithm leading to reduced implementation complexity compared to that of [93]. For the method presented in [95], the input signal needs to have certain spectrum sparsity to ensure that there exist some alias-free frequency points in the output spectrum. The phase information of the output spectra of the sub-ADCs at the alias-free frequency points is exploited to estimate the time skews. The estimation accuracy improves with higher sparsity.

The methods in [27, 96-99] are based on the observation that the signal changes more on average if the time interval between samples is longer and vice versa. Time skews are estimated by calculating

the mean squared difference of the output samples between adjacent sub-ADCs [96-98]. The algorithm presented in [96] gives a bias in the estimate and the accuracy is affected by the uncorrected gain and offset mismatches [97]. Further, the algorithm only works well for signals bandlimited to about 1/3 of the Nyquist frequency. The time skews were estimated by minimizing the cross-correlation functions between the adjacent channels using a stochastic gradient searching [27, 99] or a steepest descent algorithm [100]. With a wide sense stationary (WSS) input signal, the output resulted from an ideal ADC is WSS and wide sense cyclostationary (WSCS) from a nonideal TIADC. In [101-102], time skews were found iteratively by minimizing a cost function parameterized on the estimates such that the WSS property of the reconstructed signal is restored gradually. The methods in [27, 99] are computationally very demanding because of the FFT and inverse FFT (IFFT) calculations of large data blocks involved in each iteration. The same can be said for the methods in [101-102] because the synthesis filter bank needs to be redesigned for each new estimate. Therefore, real time implementation is limited to relatively low sampling rate and small number of channels.

The authors of [22] and [103] have demonstrated a chopper based time skew estimation technique for a two-channel TIADC. Chopping the TIADC output and delaying its phase by $\pi/2$ causes the signal and image to switch position in the spectrum. The product of the chopped and un-chopped output accumulated over time produces a DC term proportional to the timing mismatch between the two channels. Estimation and correction are performed simultaneously using an adaptive LMS algorithm. In the four-channel case presented in [28], Channel-1 output is combined with the output of Channel-0 and Channel-2 separately to obtain two sequences which, after being correlated with their respective chopped output, are summed and averaged over time to produce the time skew of Channel-1. The time skews of Channel-2 and Channel-3 with reference to Channel-0 and Channel-1, respectively, are then estimated using the method in [103]. The estimation takes longer time as the convergence of the time skews in different channels are achieved in order. Further, the number of interleaving channels is limited to radix of two. A related correlation based estimation method which employs the Hadamard transform has been reported in [29]. The aliasing signals obtained by applying the Hadamard transform to the TIADC output, after being differentiated, are correlated with the original sequence to estimate the time skews in all channels at the same time. Extension to arbitrary number of channels is

straightforward by extending the Hadamard transform. The methods in [22, 28-29, 103] have a common limitation of requiring frequency null at $kf_s/2M, k = 0, 1, \dots, M - 1$.

2.3.4 Frequency Response Mismatch

Frequency response mismatch introduces additional frequency-dependent gain and timing mismatches. Frequency-dependent mismatch correction is important for achieving high-resolution over a wide frequency range and has been studied in various research works [104-126]. Among these compensation techniques, [104-107] correct for the bandwidth mismatch whereas [108-126] deal with the more general frequency response mismatch. One way to avoid performance degradation caused by bandwidth mismatch is to design S/H circuits with bandwidths which are much higher than the maximum input frequency. However, high bandwidth places great burden on the circuit design and cannot suppress high frequency noise. Hence, digital correction is preferable.

In [104], the reconstruction filter bank is designed based on the analysis of the *signal function* and *aliasing function* for a single tone sinusoidal input. Bandwidths of the S/H circuits need to be known *a priori* to design the reconstruction filters. In [105], the design complexity for a two-channel TIADC is reduced to a single FIR filter by using a first order approximation of the aliasing component matrix of the analysis filter bank and further rearrangement. The semi-blind background scheme for estimating the relative bandwidth mismatch in a two-channel TIADC requires the input signal has no spectral content near DC [106]. By utilizing a stochastic gradient algorithm, the estimate of the relative bandwidth mismatch is updated such that the image spectra of the two channels at the low pass band are approximated as much as possible. The correction structure is similar to that in [105] except that it is also adaptive to the variations in bandwidth mismatch. Generalization to arbitrary number of channels was not addressed. In [107], bandwidth mismatch in a communication receiver where the transmitted data has a white spectrum is corrected by using an adaptive M -PTV feed-forward equalizer and the coefficients of the equalizer are updated by using an LMS loop.

In [108], the corrected output spectrum is obtained by multiplying the output spectrum of each sub-

ADC with a frequency-dependent correction term and summing the results together. In frequency domain, the sub-ADC outputs and the frequency shifted copies of the input are related by a linear system. Reconstruction in [109] is achieved by computing the inverse of the $M \times M$ mismatch matrix. The algorithms in [108-109] involve extensive calculations, i.e., DFTs of the sub-ADC outputs and complex matrix inversion, and are therefore not suitable for real time processing. The method presented in [110] compensates for the general frequency response mismatch using the multirate filter bank approach where the *signal* and *aliasing functions* on a dense grid of frequency points spanning over the band of interest are controlled simultaneously. The optimization is performed using a WLS framework. The method requires the inversion of a matrix of size $MN \times MN$ with N being the length of each FIR filter. The multichannel filtering approach in [111] is to estimate the value of an ideal sample using a weighted combination of the raw output samples from all the sub-ADCs sampling in the neighborhood of the sample under estimation. The difference between the ideal and estimated frequency responses is then evaluated on a dense frequency grid and minimized using a WLS approach. Solving the problems in [111] requires the inversion of M separate matrices, each with size $N \times N$, where N is the filter length. It was reported in [111] that for the same filter length, the multichannel filtering approach requires less computational effort and is numerically more robust. A straightforward extension from the timing mismatch correction in [59] to the more general frequency response mismatch correction was presented in [112]. The identification of the channel frequency response mismatch was however not mentioned. If the correction filters of [110], [111] and [112] are designed using the same WLS approach with the same design parameters, the resulting filters are equivalent. Similar to the time skew correction technique in [64], magnitude response mismatch is corrected by first reproducing the image spectrum and then subtracting it from the TIADC output [113]-[114]. Implementation of the correction structure requires $(M - 1)$ FIR filters and $(M - 1)$ time-varying multipliers switching at the rate of f_s . Phase response mismatch, however, was not treated. The compensation structures in [64-65, 113-114] have been generalized to a unified model [115]. The authors demonstrated that the error signal arise from the general frequency response mismatch can be reproduced by using the output samples and proper filtering. The model can be implemented in various forms and achieve compensation performance as good as desired.

Exact models describing the channel transfer function is usually not available. In [110-111], the need for an accurate model was avoided by characterizing the responses of the sub-ADCs at discrete frequency points. Therefore, a characterizing stage is necessary at system start up or when the frequency responses vary. Polynomial models have been adopted in a number of research works [116-120]. In [116], the channel frequency responses are expressed in terms of polynomials in $j\omega$ where ω is the discrete-time frequency variable. The resulting compensation structure has a cascaded form where the residue error is reduced successively and becomes higher power terms of $j\omega$ with diminishing coefficients. The structure consists of a number of fixed digital filters approximating different orders of $j\omega$, and a few variable multipliers corresponding to the coefficients in the polynomial model. The values of the variable multipliers need to be re-calculated whenever the frequency responses vary. Another technique which approximates the channel frequency responses of a TIADC with a PTV FIR filter and treats the reconstruction as a linear inverse problem of a time-varying system shown in (2.26) was demonstrated in [117]-[118]. A sample-by-sample correction framework employing Gauss-Seidel iteration was presented. Farrow-based filters were adopted to cope with the variations in the channel frequency response. Farrow filter implementation is available when the mismatch only depends on a single parameter such as the time skew or bandwidth. When multiple sources of mismatch are present, cascading of the correction structures to deal with different types of mismatches is necessary. Extension of the gain and timing mismatch correction in [101]-[102] to the frequency response mismatch for a two-channel TIADC was demonstrated in [119]-[120]. The magnitude and phase responses are modeled by polynomials in ω and the polynomial parameters are estimated iteratively by minimizing the norm of a cyclic correlation function of the signal reconstructed using the previous estimates [119]. An analytic approach which avoids matrix inversion and time-frequency transformation and is thus numerically less intensive was reported in [120].

In [121], the coefficients of the time-varying compensation filter are updated using an M -periodic LMS algorithm so that the difference between the TIADC output and the output of a low-resolution ADC is minimized. The approach requires an extra reference ADC working at the overall sampling rate. The method in [122] corrects for the frequency response mismatch in a two-channel TIADC and is an extension of the work in [94]. The frequency dependent scaling function $A_1(j\Omega)/A_0(j\Omega)$ is

approximated by a polynomial function of $j\omega$ and a filtered-error LMS algorithm is used to update the estimates of the polynomial parameters by minimizing the difference between the residue error and the reproduced error in the signal free region. Adaptive filtering may also be found in [123-126]. The estimation methods impose certain constraints on the input signal. For example, in [123], the input signal is bandlimited and WSS. For the correlation based method in [125], the input signal cannot have spectral content at frequencies $f_s/8$, $f_s/4$, $3f_s/8$ etc. With the estimated mismatch parameters, correction is then achieved by using LPTV FIR filters [124-126] or FIR filter banks [123].

2.4 Chapter Summary

In this chapter, we reviewed the architecture of a TIADC as an effective means to achieve high sampling speed without place too much burden on the device technology. Unfortunately, fabrication imperfections give rise to mismatches among the sub-ADC channels which lead to deteriorated TIADC performance. In particular, four types of mismatch errors, namely, offset, gain, timing and frequency response mismatches were studied and their impacts in both the time and frequency domains were analyzed. After that, research work in over twenty years for estimating and correcting the mismatch errors was reviewed.

Among the various correction methods, digital techniques attracted most research attention because of their superior accuracy, great reliability and independence of the process technology. The ideal samples are reconstructed or interpolated using digital reconstruction system which usually employs one to several FIR filters. The resources on the FPGAs and DSPs required for realizing the digital filters and the computational burden very much depend on the filter complexity. Reducing the filter complexity, while maintaining the desired compensation performance, is essential to the design of real-time digital compensation systems. In the subsequent chapters, we shall present a new compensation structure consisting of recursive filters which has reduced complexity than the conventional FIR compensation filters.

CHAPTER 3 : DESIGN PRINCIPLE OF THE IIR FILTERS FOR TIADC MISMATCH COMPENSATION

3.1 Introduction

In this chapter, we shall present the fundamental design principle of the IIR filters for compensating the timing and frequency response mismatches of the TIADC. The coefficients of the proposed IIR filter are optimized using a WLS approach such that the difference between the reconstructed and the desired frequency response is minimized.

3.2 The Proposed IIR Mismatch Compensation Filter

Recall that if a continuous-time signal $x_a(t)$ is sampled by an ideal ADC at a rate of $f_s = 1/T_s$, ignoring the quantization error, then the resulted ideally sampled sequence $\check{x}[n]$ is

$$\check{x}[n] = x_a(nT_s), \quad n \in \mathbb{Z} \quad (3.1)$$

where $\mathbb{Z}$ represents the space containing all integers. Assume $x_a(t)$ is bandlimited according to $\Omega_b \leq \pi f_s$ so that the Nyquist criterion is fulfilled, $x_a(t)$ can be uniquely recovered from $\check{x}[n]$ by considering the Fourier transform of $\check{x}[n]$ over the frequency band $[-\pi f_s, \pi f_s]$, i.e.,

$$\dot{X}(e^{j\Omega T_s}) = \frac{1}{T_s} X_a(j\Omega), \quad |\Omega| \leq \pi f_s \quad (3.2)$$

where $\dot{X}(e^{j\Omega T_s})$ and $X_a(j\Omega)$ denote the Fourier transforms of the ideally sampled sequence $\dot{x}[n]$ and the continuous-time signal $x_a(t)$, respectively.

Figure 3.1 A hybrid analog-digital system modelling an M -channel TIADC.

However, if the continuous-time signal $x_a(t)$ is sampled by a nonideal TIADC with the overall sampling rate f_s , a different sequence will be obtained. We denote this nonideal sequence as $x[n]$. Figure 3.1 shows a hybrid analog-digital system which models a nonideal TIADC consisting of M parallel sub-ADCs. The analog transfer function $G_m(j\Omega)$, $m = 0, 1, \dots, M - 1$ represents the gain, timing and frequency response mismatches in Channel- m . Offset mismatch in Channel- m is equivalent to adding a DC voltage to the input of the m -th sub-ADC and therefore cannot be modelled by $G_m(j\Omega)$. Correction of the offset mismatch error can be done fairly easily by attaching a calibration voltage at the input of the respective sub-ADC or subtracting an equivalent digital code from the output sequence of the respective channel. Throughout this chapter and the chapters follow, the TIADC is assumed to have no offset mismatch and we aim to compensate for the mismatch errors modelled by $G_m(j\Omega)$, $m = 0, 1, \dots, M - 1$. The desired sampling time displacements between the channels are modelled by the analog functions $1, e^{j\Omega T_s}, \dots, e^{j(M-1)\Omega T_s}$ where T_s is ideal sampling interval of the TIADC, i.e., $T_s = 1/f_s$. After absorbing all the mismatches into $G_m(j\Omega)$, the M sub-ADCs are assumed to be free from mismatch errors and sample at the same time instant. Denote the

analog signal at the output of $G_m(j\Omega)$ as $x_{G,m}(t)$. $x_{G,m}(t)$ is then sampled at the channel sampling rate f_s/M to produce the raw output sequence of ADC- m , $x_m[k]$. The final output sequence of the nonideal TIADC, $x[n]$, is formed by interleaving the M sub-sequences $x_m[k]$, $m = 0, 1, \dots, M - 1$, i.e.,

$$x[n] = x_{\langle n \rangle M} \left(\left\lfloor \frac{n}{M} \right\rfloor \right), \quad (3.3)$$

where $\langle n \rangle M$ denote the n modulo M operation and $\lfloor a \rfloor$ is the maximum integer smaller than or equal to a . Reversely, we have

$$x_m[k] = x[kM + m]. \quad (3.4)$$

Note that the different time indices k and n of $x_m[k]$ and $x[n]$ signify that $x_m[k]$ and $x[n]$ are clocked at different rates, i.e., $x_m[k]$ at f_s/M and $x[n]$ at f_s . The notations used to denote the various signals appearing in this chapter are listed in Table 3.1 for ease of reference. The relationships between the ideal ADC output sample $\check{x}[n]$, the nonideal TIADC output sample $x[n]$ and the sub-ADC output sample $x_m[k]$ are demonstrated using a three-channel TIADC example shown in Figure 3.2.

Table 3.1 List of the notations for the various signals appearing in Chapter 3

$\check{x}[n]$:	The n -th output sample produced by an ideal ADC.
$x[n]$:	The n -th output sample produced by a nonideal (with mismatch) TIADC clocked at f_s .
$\hat{x}[n]$:	The estimated value of $\check{x}[n]$.
$y[n]$:	The estimated value of the ideal sample that is η samples delayed from $\check{x}[n]$, i.e., $y[n] = \hat{x}[n - \eta]$.
$x_m[k]$:	The k -th output sample of the m -th sub-ADC clocked at f_s/M .
$\hat{x}_m[k]$:	The estimate of the ideal value of $x_m[k]$.

In the multichannel FIR filtering approach demonstrated in [111], the value of the ideal sample $\check{x}[n]$ is estimated by using a weighted combination of the raw output samples from all the sub-ADCs sampled consecutively in the neighborhood of $\check{x}[n]$. The weight of each sample corresponds to the filter coefficient of the FIR filters in the multichannel network for mismatch compensation. Take the three-

channel TIADC shown in Figure 3.2(b) for example, the ideal value of $x_1[1]$, or $\ddot{x}[4]$, may be estimated using a weighted combination of the seven consecutive samples, $x_1[2]$, $x_0[2]$, $x_2[1]$, $x_1[1]$, $x_0[1]$, $x_2[0]$ and $x_1[0]$. Let $\hat{x}_m[k]$ denote the estimate of the ideal value of $x_m[k]$ and we may write $\hat{x}_1[1] = a_{1,1}[-1]x_1[2] + a_{1,0}[-1]x_0[2] + a_{1,2}[0]x_2[1] + a_{1,1}[0]x_1[1] + a_{1,0}[0]x_0[1] + a_{1,2}[1]x_2[0] + a_{1,1}[1]x_1[0]$. In general, the estimated value, $\hat{x}_m[k]$, may be expressed as

$$\begin{aligned} \hat{x}_m[k] &= \cdots + a_{m,1}[-1]x_1[k+1] + a_{m,0}[-1]x_0[k+1] + a_{m,M-1}[0]x_{M-1}[k] + \cdots \\ &\quad + a_{m,0}[0]x_0[k] + a_{m,M-1}[1]x_{M-1}[k-1] + \cdots \\ &= \sum_{i,p} a_{m,i}[p] x_i[k-p], \end{aligned} \quad (3.5)$$

where $a_{m,i}[p]$ is the p -th partial coefficients in estimating the ideal output of ADC- m using the output from ADC- i and the summation is performed over the interested range of $x_i[k-p]$.

Figure 3.2 An example of a three-channel TIADC demonstrating the relationships between the ideal ADC output sample $\ddot{x}[n]$, the sub-ADC output sample $x_m[k]$, the compensated sub-ADC output sample $\hat{x}_m[k]$, the nonideal TIADC output sample $x[n]$, and the compensated value for $x[n-3]$, $y[n]$.

Suppose there is a set of samples which are closer to the ideal samples than the TIADC raw outputs and this set of samples can be incorporated in the estimation process. Intuitively, the estimation result would be more accurate if the same number of samples is used for estimation or fewer samples would

be needed to reach the same desired accuracy. The outputs after mismatch correction, i.e., $\hat{x}_m[k]$, $m = 0, 1, \dots, M-1$, are closer to the ideal samples than the unprocessed TIADC outputs. In general, at the time of estimating the ideal value of $x_m[k]$, the values of $\hat{x}_i[k]$ where $i < m$ and $\hat{x}_i[l]$ where $i = 0, 1, \dots, M-1, l < k$ are already available and may be incorporated in the estimation process. Therefore, the value of $\hat{x}_m[k]$ may be obtained by using a weighted combination of the raw outputs from all the sub-ADCs sampled consecutively in the neighborhood of $x_m[k]$ as well as the past estimated outputs which are closest to $x_m[k]$, i.e.,

$$\begin{aligned} \hat{x}_m[k] &= \dots + a_{m,1}[-1]x_1[k+1] + a_{m,0}[-1]x_0[k+1] + a_{m,M-1}[0]x_{M-1}[k] + \dots \\ &\quad + a_{m,0}[0]x_0[k] + a_{m,M-1}[1]x_{M-1}[k-1] + \dots - b_{m,m-1}[0]\hat{x}_{m-1}[k] - \dots \\ &\quad - b_{m,M-1}[1]\hat{x}_{M-1}[k-1] - \dots \\ &= \sum_{i,p} a_{m,i}[p] x_i[k-p] - \sum_{i,q} b_{m,i}[q] \hat{x}_i[k-q]. \end{aligned} \quad (3.6)$$

In (3.6), $\hat{x}_m[k]$ is expressed as a weighted combination of $x_i[k-p]$ and $\hat{x}_i[k-q]$ where the coefficients $a_{m,i}[p]$ for all i and p are the fractional contribution of $x_i[k-p]$ to $\hat{x}_m[k]$ and the coefficients $-b_{m,i}[q]$ for all i and q are the fractional contribution of $\hat{x}_i[k-q]$ to $\hat{x}_m[k]$. The estimation structure in (3.6) is in the form of a multichannel IIR filter with M input ports. The weights of the raw samples and the past estimated output samples correspond to, respectively, the feed-forward and the feedback coefficients of the IIR filter. There are a total of M such IIR filters in the multichannel mismatch compensation network. In the three-channel TIADC example shown in Figure 3.2(b), the values of $\hat{x}_2[0]$ and $\hat{x}_0[1]$ are available at the time of computing the value of $\hat{x}_1[1]$ and we may write

$$\begin{aligned} \hat{x}_1[1] &= a_{1,1}[-1]x_1[2] + a_{1,0}[-1]x_0[2] + a_{1,2}[0]x_2[1] + a_{1,1}[0]x_1[1] + a_{1,0}[0]x_0[1] \\ &\quad + a_{1,2}[1]x_2[0] + a_{1,1}[1]x_1[0] - b_{1,0}[0]\hat{x}_0[1] - b_{1,2}[1]\hat{x}_2[0]. \end{aligned} \quad (3.7)$$

The filter for computing the value of $\hat{x}_1[k]$ is a multichannel IIR filter with three input ports and its implementation is shown in Figure 3.3. The values of $\hat{x}_0[k]$ and $\hat{x}_2[k]$ are computed using different sets of filter coefficients which are not shown in the figure. Note that $\hat{x}_{m+1}[k]$, $\hat{x}_m[k]$, and $\hat{x}_{m-1}[k]$ are consecutive TIADC output samples but are computed using different filters whereas $\hat{x}_m[k+1]$,

$\hat{x}_m[k]$, and $\hat{x}_m[k-1]$ are M samples apart and are computed using the same filter. Figure 3.4 shows the general structure of a multichannel IIR filter for estimating the ideal value of $x_m[k]$ in an M -channel TIADC. A suitable integer delay has been added to the system to ensure causality. Note that $b_{m,i}[0]$ exists for $i < m$ since $\hat{x}_i[k]$ is already available at the time of computing $\hat{x}_m[k]$. The multichannel FIR filter presented in [111] may be regarded as a special class of the proposed IIR filter with zero-th order denominator in the filter's transfer function.

Figure 3.3 Multichannel IIR filter structure implementation for computing $\hat{x}_1[k]$.

Figure 3.4 The general multichannel IIR filter structure with M channels for estimating the low rate sequence $\hat{x}_m[k]$.

The reconstruction system described in (3.6) consists of M multichannel IIR filters which operate on the low rate sub-sequences from all the sub-ADCs. In fact, the set of M multichannel filters may equivalently be interpreted as a linear M -periodic time-varying (M -PTV) IIR filter operating on the high rate sequence $x[n]$. Denoting $\hat{x}[n]$ as the estimated value of $\check{x}[n]$ and using a different notation, the estimation equation in (3.6) may be rewritten in terms of the high rate data sequences as

$$\hat{x}[n] = \sum_p a_n[p]x[n-p] - \sum_q b_n[q]\hat{x}[n-q]. \quad (3.8)$$

Using the above notations, (3.7) can be rewritten as $\hat{x}[4] = \sum_{p=-3}^3 a_4[p]x[4-p] - \sum_{q=1}^2 b_4[q]\hat{x}[4-q]$. Obviously, $a_4[-3] = a_{1,1}[-1]$, $a_4[-2] = a_{1,0}[-1]$, $b_4[1] = b_{1,0}[0]$ etc. In (3.8), $\hat{x}[n]$ is generated through a time-varying IIR filter and the subscript n of $a_n[p]$ and $b_n[q]$ indicates the time dependence of the IIR filter. The filter coefficients $a_n[p]$ and $b_n[q]$ are also periodic with period M , i.e., $a_n[p] = a_{n+M}[p]$ and $b_n[q] = b_{n+M}[q]$. Therefore, any two output codes spaced M samples apart are estimated using the same set of filter coefficients. The estimation process described in (3.8) becomes noncausal when $p < 0$. To obtain a causal reconstruction system, a suitable integer delay, η , can be introduced so that we aim to estimate the past output of the TIADC rather than its immediate output. η represents the group delay introduced by the compensation filter and the choice of η affects the stability of the proposed IIR compensation filter. The details will be discussed in Chapter 5. Let $y[n]$ denote the estimated value of $\check{x}[n - \eta]$ and it may be formed as

$$y[n] = \sum_{p=0}^{N_a-1} a_n[p]x[n-p] - \sum_{q=1}^{N_b} b_n[q]y[n-q] \quad (3.9)$$

where N_a and N_b are the number of feed-forward and feedback coefficients, respectively, of the IIR filter for estimating one channel output. In the example demonstrated in Figure 3.2(c), a delay of $\eta = 3$ is introduced to make the system causal.

Taking z -transform of (3.9) and rearranging, we have

$$Y(z) = H_n(z)X(z), \quad (3.10)$$

where $Y(z)$ and $X(z)$ denote the z -transform of $y[n]$ and $x[n]$, respectively and $H_n(z)$ is the z -transform of the time-varying IIR mismatch compensation filter at time instant n . $H_n(z)$ is given by

$$H_n(z) = \frac{\sum_{p=0}^{N_a-1} a_n[p]z^{-p}}{1 + \sum_{q=1}^{N_b} b_n[q]z^{-q}} \quad (3.11)$$

$H_n(z)$ is also M -periodic, i.e., $H_n(z) = H_{n+M}(z)$. Figure 3.5(a) shows the block diagram of a general time-varying filter. Figure 3.5(b) shows the direct form implementation of the proposed M -PTV IIR mismatch compensation filter. The filter coefficients $a_n[p]$ and $b_n[q]$ are time-varying and periodic with period M .

Figure 3.5 (a) A general time-varying filter. (b) Direct form I implementation of the proposed time-varying IIR filter for TIADC mismatch compensation.

Let $e[n]$ denote the reconstruction error which is defined as the difference between the reconstructed output sequence $y[n]$, and the delayed ideal sequence $\check{x}[n - \eta]$. We then have

$$e[n] = y[n] - \check{x}[n - \eta]. \quad (3.12)$$

Using (3.1), (3.12) is further written as

$$e[n] = y[n] - x_a((n - \eta)T_s). \quad (3.13)$$

Ignoring the quantization error, it is easy to see from Figure 3.1 that the TIADC output sequence $x[n]$ is related to $x_{G,m}(t)$ by

$$x[n] = x_{G,\langle n \rangle M}(nT_s). \quad (3.14)$$

(3.14) implies that a TIADC in the presence of mismatch errors is also an M -PTV system. Taking the inverse Fourier transform of $x[n]$, or $x_{G,\langle n \rangle M}(nT_s)$, we have

$$x[n] = \frac{1}{2\pi} \int_{-\pi f_s}^{\pi f_s} X_a(j\Omega) G_{\langle n \rangle M}(j\Omega) e^{jn\Omega T_s} d\Omega, \quad (3.15)$$

where $X_a(j\Omega)$ is the Fourier transform of the analog input signal $x_a(t)$. Next, we want to take the inverse Fourier transform of the reconstructed sequence $y[n]$ which is not available at this stage. However, we can make a reasonable assumption that the TIADC is well compensated such that $e[n] \approx 0$ and $y[n]$ can be approximately by $\check{x}[n - \eta]$ and

$$y[n] \approx \frac{1}{2\pi} \int_{-\pi f_s}^{\pi f_s} X_a(j\Omega) e^{-j\eta\Omega T_s} e^{jn\Omega T_s} d\Omega. \quad (3.16)$$

Substituting (3.15) and (3.16) into the right hand side of (3.9) and rearranging results

$$y[n] \approx \frac{1}{2\pi} \int_{-\pi f_s}^{\pi f_s} X_a(j\Omega) T_n(j\Omega) e^{jn\Omega T_s} d\Omega \quad (3.17)$$

where

$$T_n(j\Omega) = \sum_{p=0}^{N_a-1} a_n[p] G_{(n-p)M}(j\Omega) e^{-jp\Omega T_s} - \sum_{q=1}^{N_b} b_n[q] e^{-j(\eta+q)\Omega T_s}. \quad (3.18)$$

It is easy to see that $T_n(j\Omega)$ is an M -PTV function. Substituting (3.17) into (3.13), and using the inverse Fourier transform of $x_a((n-\eta)T_s)$, we can then express the reconstruction error as

$$e[n] \approx \frac{1}{2\pi} \int_{-\pi f_s}^{\pi f_s} X_a(j\Omega) (T_n(j\Omega) - e^{-j\eta\Omega T_s}) e^{jn\Omega T_s} d\Omega. \quad (3.19)$$

Our goal is to minimize the reconstruction error $e[n]$. Since $T_n(j\Omega)$ is M -PTV, the design problem becomes one of finding the filter coefficients $a_m[p]$ and $b_m[q]$ such that $(T_m(j\Omega) - e^{-j\eta\Omega T_s})$ is minimized over the frequency range of interest, where $m = 0, 1, \dots, M-1$.

3.3 WLS Optimization Approach

WLS design of linear time invariant (LTI) IIR digital filters can be found in a number of literatures [127]-[130]. Stable LTI IIR filter designs were achieved by incorporating various stability constraints in the WLS optimization process, such as the positive realness constraint in [128], the *Rouche's theorem* based constraint in [129], the argument principle based constraint in [130] etc. This section only presents the WLS problem formulation for determining the filter coefficients of the IIR compensation filter. Methods for achieving the filter's stability will be covered in Chapter 5.

Define a new error measure $\varepsilon_m(j\Omega)$ for Channel- m as

$$\varepsilon_m(j\Omega) = \sqrt{w_m(\Omega)} (T_m(j\Omega) - e^{-j\eta\Omega T_s}), \quad (3.20)$$

where $\sqrt{w_m(\Omega)}$ is the WLS error weighting function and can be used to determine the frequency range of interest in the optimization. For example, a transformer placed at the ADC input introduces a lower cut off frequency Ω_L and the aliasing terms in the low pass band (from DC to Ω_L) is of no interest to the filter design, we may let $w_m(\Omega) = 1$ for $\Omega \geq \Omega_L$ and $w_m(\Omega) = 0$ for $0 \leq \Omega < \Omega_L$. It has been demonstrated in a number of literatures that it is advantageous to have slight oversampling

when designing the mismatch compensation filters because the filter order increases dramatically as the input frequency approaching $\Omega = \pi f_s$. In this case, we may let $w_m(\Omega) = 1$ for $0 \leq \Omega \leq \Omega_H$ and $w_m(\Omega) = 0$ for $\Omega > \Omega_H$, where Ω_H is the higher cut off frequency. Denote the coefficient vector for estimating Channel- m output as $\mathbf{h}_m$ and $\mathbf{h}_m$ is defined as

$$\mathbf{h}_m = [a_m[0], \dots, a_m[N_a - 1], b_m[1], \dots, b_m[N_b]]^T. \quad (3.21)$$

Let the total number of coefficients to be optimized be N and $N = N_a + N_b$. $\mathbf{h}_m$ is an $N \times 1$ vector.

Define the signal vector for Channel- m , $\mathbf{s}_m(j\Omega)$, as

$$\mathbf{s}_m(j\Omega) = [G_m(j\Omega), \dots, G_{(m-(N_a-1))M}(j\Omega)e^{-j(N_a-1)\Omega T_s}, -e^{-j(\eta+1)\Omega T_s}, \dots, -e^{-j(\eta+N_b)\Omega T_s}]^T. \quad (3.22)$$

With (3.21) and (3.22), we can express (3.20) in vector form as

$$\varepsilon_m(j\Omega) = \sqrt{w_m(\Omega)}([\mathbf{s}_m(j\Omega)]^T \mathbf{h}_m - e^{-j\eta\Omega T_s}). \quad (3.23)$$

Perfect reconstruction is achieved if $[\mathbf{s}_m(j\Omega)]^T \mathbf{h}_m = e^{-j\eta\Omega T_s}$ for $m = 0, 1, \dots, M-1$ and $\Omega \in [0, \pi f_s]$, in which case $y[n] = \tilde{x}[n - \eta]$. However, it is impossible to achieve perfect reconstruction by using practical compensation filters. Hence, the design goal is to find M LTI IIR filters such that $\varepsilon_m(j\Omega)$ is minimized for $m = 0, 1, \dots, M-1$ over the frequency range of interest. The compensated TIADC output using practical filters is still M -PTV, but the time dependence becomes much weaker. In other words, the distortions caused by the mismatch errors are alleviated significantly.

The coefficient vector $\mathbf{h}_m$ may be optimized by minimizing $|\varepsilon_m(j\Omega)|$ in the least squares sense over the frequency range of interest. In our approach, (3.23) is evaluated on a dense grid of frequency points $\Omega_l \in [0, \pi f_s]$, $l = 1, 2, \dots, L$. The optimization problem becomes one of minimizing the following

$$\sum_{l=1}^L |\varepsilon_m(j\Omega_l)|^2. \quad (3.24)$$

Define the complex column vectors $\boldsymbol{\varepsilon}_m$ and $\mathbf{u}_m$, the complex matrix $\mathbf{S}_m$, and the real diagonal matrix $\mathbf{W}_m$ as in (3.25) through (3.28),

$$\boldsymbol{\varepsilon}_m = [\varepsilon_m(j\Omega_1), \varepsilon_m(j\Omega_2), \dots, \varepsilon_m(j\Omega_L)]^T, \quad (3.25)$$

$$\mathbf{u}_m = [e^{-j\eta\Omega_1 T_s}, e^{-j\eta\Omega_2 T_s}, \dots, e^{-j\eta\Omega_L T_s}]^T, \quad (3.26)$$

$$\mathbf{S}_m = [\mathbf{s}_m(j\Omega_1), \mathbf{s}_m(j\Omega_2), \dots, \mathbf{s}_m(j\Omega_L)]^T, \quad (3.27)$$

$$\mathbf{W}_m = \begin{bmatrix} w_m(\Omega_1) & 0 & \dots & 0 \\ 0 & w_m(\Omega_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_m(\Omega_L) \end{bmatrix}, \quad (3.28)$$

where $\boldsymbol{\varepsilon}_m$ and $\mathbf{u}_m$ are $L \times 1$ vectors, $\mathbf{W}_m$ is an $L \times L$ diagonal matrix and $\mathbf{S}_m$ is an $L \times N$ matrix.

Using the notations in (3.25) through (3.28), we have

$$\boldsymbol{\varepsilon}_m = \sqrt{\mathbf{W}_m}(\mathbf{S}_m \mathbf{h}_m - \mathbf{u}_m). \quad (3.29)$$

The WLS optimum solution of $\mathbf{h}_m$ that minimizes $\sum_{l=1}^L |\varepsilon_m(j\Omega_l)|^2 = [\boldsymbol{\varepsilon}_m]^H \boldsymbol{\varepsilon}_m$ is given by [131] as

$$\mathbf{h}_m = (\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)^{-1} (\mathbf{S}_m^H \sqrt{\mathbf{W}_m} \mathbf{u}_m) \quad (3.30)$$

where $\mathbf{S}_m^H$ denotes the complex conjugate transpose of $\mathbf{S}_m$. The overall compensation system is solved through M separate matrix inversions, where each matrix is of size $N \times N$.

3.4 Discussion

a) Issue in forming the signal matrix $\mathbf{S}_m$

As shown in Figure 3.1, each sub-ADC channel has a channel frequency response $G_m(j\Omega)$. Mismatch occurs when at least one channel frequency response is different from the others. In many cases, it is convenient to choose a reference channel, usually Channel-0, and design IIR compensation filters for Channels 1 to $M - 1$ such that all the channels experience the same frequency response $G_0(j\Omega)$ after

compensation. This common frequency response is tolerable and may be removed by using an equalizer. Therefore, we aim to compensate for the relative mismatch error in ADC- m , $m = 1, 2, \dots, M - 1$ with respect to that in ADC-0 instead of the absolute mismatch error in every channel. ADC-0 is then assumed to be ideal and free from all mismatches, i.e., $G_0(j\Omega) = 1$. In that case, the filter for estimating ADC-0 output is just an ideal delay $e^{-j\eta\Omega T_s}$ and only $(M - 1)$ matrix inversions are needed.

The ideal sample in each channel is estimated using the raw output samples as well as the past estimated outputs from all the sub-ADCs. In obtaining the filter coefficients, the estimated output samples are not available and the ideal samples are used instead in the WLS formulation. Therefore, when forming the matrix $\mathbf{S}_m$ in (3.27), some columns correspond to the raw data of ADC-0, $x_0[k]$, and the estimated data of ADC-0, $\hat{x}_0[k]$, may be dependent and $\mathbf{S}_m$ becomes rank-deficient. For example, in (3.7), both $x_0[1]$ and $\hat{x}_0[1]$ are used to calculate $\hat{x}_1[1]$. With $\eta = 3$, the matrix $\mathbf{S}_1$ is formed by evaluating the row vector $[G_1(j\Omega), G_0(j\Omega)e^{-j\Omega T_s}, G_2(j\Omega)e^{-j2\Omega T_s}, G_1(j\Omega)e^{-j3\Omega T_s}, G_0(j\Omega)e^{-j4\Omega T_s}, G_2(j\Omega)e^{-j5\Omega T_s}, G_1(j\Omega)e^{-j6\Omega T_s}, -e^{-j4\Omega T_s}, -e^{-j5\Omega T_s}]$ at a dense frequency grid. We can see that the column corresponding to $x[3]$ or $x_0[1]$ is $G_0(j\Omega)e^{-j4\Omega T_s}$ being evaluated at different frequency points and the column corresponding to $y[6]$ or $\hat{x}_0[1]$ is $-e^{-j4\Omega T_s}$ being evaluated at different frequency points. Since $G_0(j\Omega)$ is assumed to be 1, these two columns are exactly the same except for the sign making the matrix $\mathbf{S}_1$ rank-deficient. To ensure that the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ in (3.30) is well conditioned, the corresponding compensated output samples from ADC-0 should not be used and the corresponding feedback coefficients should be assigned zero value. For the q -th feedback coefficient in the M -PTV IIR filter for estimating Channel- m output, $b_m[q]$, it should be set to zero if $\langle q - m \rangle M = 0$ and $q + \eta < N_a$. Therefore, the actual number of feedback coefficients to be optimized can be smaller than N_b .

b) Residue error

In the derivation of the coefficients of the IIR compensation filters, we assumed the TIADC is perfectly compensated. As shown in (3.16)-(3.19), the ideal output sample $x_a((n - \eta)T_s)$ is used in

place of the actual estimated output sample $y[n]$ in the problem formulation. (3.19) is only an approximation and the equality holds if the left hand side is modified to $e[n] + \sum_{q=1}^{N_b} b_n[q]e[n-q]$, i.e.,

$$e[n] + \sum_{q=1}^{N_b} b_n[q]e[n-q] = \frac{1}{2\pi} \int_{-\pi f_s}^{\pi f_s} X_a(j\Omega)(T_n(j\Omega) - e^{-j\eta\Omega T_s})e^{jn\Omega T_s} d\Omega. \quad (3.31)$$

In (3.31), $e[n-q]$, $q = 1, 2, \dots, N_b$ is the residue error when we use the corrected output samples instead of the ideal samples in the estimation process. The error measure $\varepsilon_m(j\Omega)$ in (3.23) then becomes

$$\varepsilon_m(j\Omega) = \sqrt{w_m(\Omega)}([\mathbf{s}_m(j\Omega)]^T \mathbf{h}_m - e^{-j\eta\Omega T_s})/Q_m(e^{j\Omega T_s}), \quad (3.32)$$

where

$$Q_m(e^{j\Omega T_s}) = 1 + \sum_{q=1}^{N_b} b_m[q]e^{-jq\Omega T_s}. \quad (3.33)$$

If the residue errors are included, the error measure $\varepsilon_m(j\Omega)$ becomes a nonlinear function of $\mathbf{h}_m$ and the problem of minimizing $\sum_{l=1}^L |\varepsilon_m(j\Omega_l)|^2$ cannot be solved using simple optimization algorithms such as the WLS method. In order to keep the computation complexity low, we ignore the residue errors on the ground that the TIADC is well compensated. Hence, the obtained solution is a sub-optimal one. Due to the presence of the residue error, the WLS solution given in (3.30) which minimize the term $\sum_{l=1}^L |\varepsilon_m(j\Omega_l)|^2$ actually minimizes the power of the sequence $(e[n] + \sum_{q=1}^{N_b} b_n[q]e[n-q])$. In contrast, for the FIR compensation filters in [111], the WLS solution is optimal and minimizes the power of the error sequence $e[n]$ which is a direct measure of the compensation performance.

If the multichannel FIR filters in [111] and the IIR filters introduced in this chapter are designed to achieve the same level of the WLS approximation error, then the compensation performance of the IIR filters would be worse than that of the FIR filters. In the design of the IIR compensation filters, a

few more coefficients are needed to achieve lower WLS error such that the effect of the residue error is compensated. Nevertheless, for the same level of mismatch compensation performance, the total number of coefficients required by the IIR filter is still smaller than that required by the FIR filter as will be shown later in Chapter 6.

The compensation performance of the proposed IIR compensation filter is affected by both the WLS error and the feedback coefficients. An IIR filter design with smaller WLS error than another design does not necessarily have a better mismatch compensation performance than the other design. Hence, in the design of the IIR compensation filters, we shall specify the design goal in terms of the TIADC performance after mismatch compensation rather than the WLS error.

c) System convergence due to lack of samples at system startup

During system start-up, the previous output samples that are used to compute the current output sample are not available and the output of the IIR compensation filter may not converge correctly. We shall show that as long as the IIR filter is stable, the compensator output is guaranteed to converge correctly.

Let's use a simple example to illustrate this point. Assume an M -PTV IIR filter with difference equation $y[n] = x[n] - b_n[1]y[n-1] - b_n[2]y[n-2]$. Further assume the previous output samples are available at system start-up, then at $n = 0$, we have

$$y[0] = x[0] - b_0[1]y[-1] - b_0[2]y[-2]$$

and at $n = 1$, we have

$$y[1] = x[1] - b_1[1]y[0] - b_1[2]y[-1].$$

In reality, $y[-1]$ and $y[-2]$ do not exist. The output of the real system can be viewed as the sum of two components: one is the output of the IIR filter excited by $x[n]$ with all the previous output samples available and the other is the finite sequence $\{b_0[1]y[-1] + b_0[2]y[-2], b_1[2]y[-1]\}$ being

filtered by the feedback part of the IIR filter. The first component gives the correct output. If the IIR filter is stable, the second component will decay to a very small magnitude eventually and can be ignored. Therefore, the convergence of our IIR compensation filter is guaranteed as long as the IIR filter is stable.

3.5 Chapter Summary

In this chapter, we have introduced an unprecedented mismatch compensation system comprising of IIR filters and presented the design principle of the IIR compensation filters. The proposed compensation filter can be interpreted as either multichannel IIR filters with M input ports or an M -PTV IIR filter. We have introduced a WLS optimization approach to obtain the filter coefficients. The issues which are exclusive to the design of the IIR compensation filters were also discussed.

CHAPTER 4 : STABILITY CRITERION OF THE PROPOSED IIR MISMATCH COMPENSATION FILTER

4.1 Introduction

The proposed TIADC mismatch compensation system is an M -PTV IIR filter and the feedback from the past estimated output is an essential part. Therefore the proposed IIR compensation filter can be potentially unstable. At each particular time instant, the M -PTV IIR filter can be viewed as an LTI IIR filter and there are a total of M LTI IIR filters. When estimating the output of ADC- m , the feedback is from not only the past estimated output of ADC- m , but also the past estimated output of the rest of the sub-ADCs. The M LTI IIR filters are therefore coupled together and the stability of each individual LTI IIR filter does not guarantee the stability of the entire compensation system.

It was demonstrated that by applying block processing, an M -PTV system operating at the rate of f_s can be represented by an equivalent MIMO system operating at the rate of f_s/M [132-133]. The input and output of the MIMO system are M -dimensional vectors and the MIMO system itself is LTI. The stability of the M -PTV filter is then determined by the pole locations of the transfer function matrix of the MIMO system, $\mathbf{H}(z)$. By using a state space representation of the MIMO system, it is further shown that the proposed system is stable if the eigenvalues of the *state transition matrix*, $\tilde{\mathbf{A}}$, have magnitudes smaller than one. Utilizing the equivalent MIMO model, the M -PTV IIR filter may alternatively be realized as a synthesis filter bank consisting of M parallel LTI IIR filters sharing a common denominator [134]. In this chapter, we shall show the detailed derivations of the equivalent

MIMO and filter bank representations. After that, an alternative approach to derive the transfer function matrix of the M -PTV IIR filter without using the state space representation is presented.

Although the stability criterion of the M -PTV IIR filter can be derived explicitly, it is difficult to incorporate the stability constraint into the WLS optimization process. We shall introduce two simple “computer-aided” approaches for gradually achieving a stable IIR filter design in Chapter 5. The eigenvalue based stability criterion is merely used to check whether a stable design has been reached in the design process.

4.2 Equivalent MIMO Representation of a PTV System

4.2.1 MIMO Representation of an M -PTV IIR Filter

Figure 4.1 Direct form I implementation of the proposed M -PTV IIR compensation filter with the state variables indicated.

We shall show how the proposed M -PTV IIR filter is transferred into an LTI MIMO system with vector input and vector output by using the state space representation. Recall that the time domain input-output relationship of the proposed IIR compensation filter is

$$y[n] = \sum_{p=0}^{N_a-1} a_n[p]x[n-p] - \sum_{q=1}^{N_b} b_n[q]y[n-q]. \quad (4.1)$$

The signal flow in a time-varying IIR filter is different from that in a time invariant one. It can be easily verified with simple examples that a time-varying IIR filter as depicted in (4.1) can only be implemented in the “Direct form I” structure as shown in Figure 4.1. In the digital filter implementation, the outputs of the delay elements are called state variables. For the implementation structure in Figure 4.1, a total of $N_s = N_a + N_b - 1$ delay elements are used and thus we have N_s state variables, $s_1[n], \dots, s_{N_b}[n], r_1[n], \dots, r_{N_a-1}[n]$, as indicated in the figure. With the state variables defined, the following set of equations can be written

$$\begin{aligned} s_1[n+1] &= s_2[n], \\ s_2[n+1] &= s_3[n], \\ &\vdots \\ s_{N_b}[n+1] &= \sum_{p=1}^{N_a-1} a_n[N_a-p]r_p[n] - \sum_{q=1}^{N_b} b_n[N_b+1-q]s_q[n] + a_n[0]x[n], \\ r_1[n+1] &= r_2[n], \\ r_2[n+1] &= r_3[n], \\ &\vdots \\ r_{N_a-1}[n+1] &= x[n], \\ y[n] &= \sum_{p=1}^{N_a-1} a_n[N_a-p]r_p[n] - \sum_{q=1}^{N_b} b_n[N_b+1-q]s_q[n] + a_n[0]x[n]. \end{aligned} \quad (4.2)$$

The set of equations in (4.2) may be rearranged and expressed in matrix-vector form as

$$\mathbf{s}[n+1] = \mathbf{A}_n \mathbf{s}[n] + \mathbf{b}_n x[n] \quad (4.3a)$$

$$y[n] = \mathbf{c}_n^T \mathbf{s}[n] + d_n x[n] \quad (4.3b)$$

where $\mathbf{s}[n]$ represents the vector of state variables at time instant n and has a dimension of $N_s \times 1$; $\mathbf{A}_n$ is an $N_s \times N_s$ matrix; $\mathbf{b}_n$ and $\mathbf{c}_n$ are both $N_s \times 1$ vectors; $d_n, x[n]$ and $y[n]$ are scalars; $\mathbf{s}[n], \mathbf{A}_n, \mathbf{b}_n,$

$\mathbf{c}_n$, and d_n are defined as in (4.4) through (4.8) as

$$\mathbf{s}[n] = [s_1[n], s_2[n], \dots, s_{N_b}[n], r_1[n], r_2[n], \dots, r_{N_a-1}[n]]^T, \quad (4.4)$$

$$\mathbf{A}_n = \begin{bmatrix} 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ -b_n[N_b] & -b_n[N_b-1] & \dots & -b_n[1] & a_n[N_a-1] & a_n[N_a-2] & \dots & a_n[1] \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad (4.5)$$

$$\mathbf{b}_n = \begin{bmatrix} \mathbf{0}_{(N_b-1) \times 1} \\ a_n[0] \\ \mathbf{0}_{(N_a-2) \times 1} \\ 1 \end{bmatrix}, \quad (4.6)$$

$$\mathbf{c}_n = \begin{bmatrix} -b_n[N_b] \\ \vdots \\ -b_n[1] \\ a_n[N_a-1] \\ \vdots \\ a_n[1] \end{bmatrix}, \quad (4.7)$$

$$d_n = a_n[0], \quad (4.8)$$

where $\mathbf{0}_{m \times n}$ denotes the m -row by n -column null matrix/vector with all its entries being zero. (4.3a-b) together with the state space parameters defined in (4.4)-(4.8) is the state space representation of the time-varying IIR filter shown in Figure 4.1. The state space representation involves the input, the state variables and the output and completely describes the system behaviour of the time-varying IIR filter in the time domain. (4.3a) and (4.3b) are called the *state equation* and the *output equation*, respectively. $\mathbf{s}[n]$ represents the memory of the system and is called the *state vector* and $\mathbf{A}_n$ is called the *state transition matrix*. For an M -PTV IIR filter, the filter coefficients and hence the state space parameters $(\mathbf{A}_n, \mathbf{b}_n, \mathbf{c}_n, d_n)$ are M -PTV as well, i.e., $\mathbf{A}_n = \mathbf{A}_{n+M}$, $\mathbf{b}_n = \mathbf{b}_{n+M}$, $\mathbf{c}_n = \mathbf{c}_{n+M}$, and $d_n = d_{n+M}$.

If the M -PTV filter is described in terms of blocks of input and output and the block length is integer multiples of M , the block representation of the M -PTV filter is time-invariant [133]. To derive the block state space representation of the proposed IIR compensation filter, let's first define a new state vector $\mathbf{t}[n]$ as the down-sampled version of $\mathbf{s}[n]$ such that $\mathbf{t}[n] = \mathbf{s}[nM]$. Using the state equation in

(4.3a) and the M -periodicity of $\mathbf{A}_n$, $\mathbf{t}[n + 1]$ can be written as

$$\begin{aligned}\mathbf{t}[n + 1] &= \mathbf{s}[nM + M] \\ &= \mathbf{A}_{M-1}\mathbf{s}[nM + M - 1] + \mathbf{b}_{M-1}x[nM + M - 1].\end{aligned}\tag{4.9}$$

Using (4.3a) to further expand $\mathbf{s}[nM + M - 1]$ in terms of $\mathbf{s}[nM + M - 2]$ and $x[nM + M - 2]$, and continuing the process until $\mathbf{s}[nM]$ is reached

$$\begin{aligned}\mathbf{t}[n + 1] &= (\mathbf{A}_{M-1}\mathbf{A}_{M-2} \dots \mathbf{A}_1\mathbf{A}_0)\mathbf{s}[nM] + (\mathbf{A}_{M-1}\mathbf{A}_{M-2} \dots \mathbf{A}_1\mathbf{b}_0)x[nM] + \dots \\ &\quad + (\mathbf{A}_{M-1}\mathbf{b}_{M-2})x[nM + M - 2] + \mathbf{b}_{M-1}x[nM + M - 1].\end{aligned}\tag{4.10}$$

Denote $\mathbf{x}[n]$ as the input signal vector consisting a block of M input samples, we have

$$\mathbf{x}[n] = [x[nM], x[nM + 1], \dots, x[nM + M - 1]]^T.\tag{4.11}$$

(4.10) can be rewritten in terms of the input signal vector as

$$\mathbf{t}[n + 1] = \tilde{\mathbf{A}}\mathbf{t}[n] + \tilde{\mathbf{B}}\mathbf{x}[n],\tag{4.12}$$

where

$$\tilde{\mathbf{A}} = \mathbf{A}_{M-1}\mathbf{A}_{M-2} \dots \mathbf{A}_1\mathbf{A}_0,\tag{4.13}$$

$$\tilde{\mathbf{B}} = [(\mathbf{A}_{M-1}\mathbf{A}_{M-2} \dots \mathbf{A}_1\mathbf{b}_0), \dots, (\mathbf{A}_{M-1}\mathbf{b}_{M-2}), \mathbf{b}_{M-1}].\tag{4.14}$$

Define the $\mathbf{y}[n]$ as the output signal vector consisting a block of M output samples, we have

$$\mathbf{y}[n] = [y[nM], y[nM + 1], \dots, y[nM + M - 1]]^T.\tag{4.15}$$

Expanding each element of $\mathbf{y}[n]$ in a similar manner and combining the equations in matrix form leads to

$$\mathbf{y}[n] = \tilde{\mathbf{C}}\mathbf{t}[n] + \tilde{\mathbf{D}}\mathbf{x}[n],\tag{4.16}$$

where

$$\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{c}_0^T \\ \mathbf{c}_1^T \mathbf{A}_0 \\ \mathbf{c}_2^T \mathbf{A}_1 \mathbf{A}_0 \\ \vdots \\ \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{A}_{M-3} \cdots \mathbf{A}_0 \end{bmatrix}, \quad (4.17)$$

and $\tilde{\mathbf{D}}$ is a lower triangular matrix with its element at i -th row and j -th column, $\tilde{d}_{ij}$, defined as, assuming the indices i and j starts from zero

$$\tilde{d}_{ij} = \begin{cases} 0 & i < j \\ d_i & i = j \\ \mathbf{c}_i^T \mathbf{b}_j & i = j + 1 \\ \mathbf{c}_i^T \mathbf{A}_{i-1} \cdots \mathbf{A}_{j+1} \mathbf{b}_j & i \geq j + 2 \end{cases} \quad (4.18)$$

The detailed derivation of the block state space representation can be found in Appendix A. Equations (4.11)-(4.18) fully describe the time domain behavior of the M -PTV IIR filter in terms of blocks of input and output where the block length is M . The four state space parameters in the block state space representation, $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\tilde{\mathbf{C}}$, and $\tilde{\mathbf{D}}$ have dimensions of $N_s \times N_s$, $N_s \times M$, $M \times N_s$, and $M \times M$, respectively. Using block processing, the proposed IIR compensation filter is now transformed into an equivalent M -input, M -output MIMO system with vector input and vector output given by

$$\mathbf{x}[k] = [x_0[k], x_1[k], \dots, x_{M-1}[k]]^T, \quad (4.19)$$

$$\mathbf{y}[k] = [y_0[k], y_1[k], \dots, y_{M-1}[k]]^T, \quad (4.20)$$

where $x_m[k]$, $m = 0, 1, \dots, M-1$ is the sub-sequence produced by the m -th sub-ADC and $y_m[k]$, $m = 0, 1, \dots, M-1$ is the compensated output of Channel- m . Figure 4.2 shows the equivalent MIMO system with the M inputs being the sub-sequences produced by the M sub-ADCs.

The MIMO system is time-invariant and operates at the rate of f_s/M . Taking the z -transform of (4.12) and (4.16) and we have

$$z\mathbf{T}(z) = \tilde{\mathbf{A}}\mathbf{T}(z) + \tilde{\mathbf{B}}\mathbf{X}(z), \quad (4.21)$$

$$\mathbf{Y}(z) = \tilde{\mathbf{C}}\mathbf{T}(z) + \tilde{\mathbf{D}}\mathbf{X}(z). \quad (4.22)$$

Figure 4.2 Equivalent MIMO system with vector input and vector output of the M -PTV IIR filter.

Rearranging (4.21) and substituting it into (4.22) leads to

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{X}(z), \quad (4.23)$$

where $\mathbf{H}(z)$ is the transfer function matrix of the equivalent MIMO system and is given by

$$\mathbf{H}(z) = \tilde{\mathbf{D}} + \tilde{\mathbf{C}}(z\mathbf{I} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{B}}, \quad (4.24)$$

where $\mathbf{I}$ denotes the identity matrix of dimension $N_s \times N_s$. $\mathbf{H}(z)$ is of size $M \times M$ and has the following form

$$\mathbf{H}(z) = \begin{bmatrix} H_{00}(z) & H_{01}(z) & \cdots & H_{0,M-1}(z) \\ H_{10}(z) & H_{11}(z) & \cdots & H_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1,0}(z) & H_{M-1,1}(z) & \cdots & H_{M-1,M-1}(z) \end{bmatrix}. \quad (4.25)$$

The entry of $\mathbf{H}(z)$ at the i -th row and j -th column, $H_{i,j}(z)$, represents the transfer function from the j -th input port to the i -th output port. (4.24) can be further written as

$$\mathbf{H}(z) = \tilde{\mathbf{D}} + \frac{\tilde{\mathbf{C}} \cdot \text{adj}(z\mathbf{I} - \tilde{\mathbf{A}}) \cdot \tilde{\mathbf{B}}}{\det(z\mathbf{I} - \tilde{\mathbf{A}})}, \quad (4.26)$$

where $\text{adj}(z\mathbf{I} - \tilde{\mathbf{A}})$ and $\det(z\mathbf{I} - \tilde{\mathbf{A}})$ denote the adjugate and the determinant of the square matrix $(z\mathbf{I} - \tilde{\mathbf{A}})$, respectively. $\det(z\mathbf{I} - \tilde{\mathbf{A}})$ is the characteristic polynomial of the matrix $\tilde{\mathbf{A}}$ and its roots are the eigenvalues of $\tilde{\mathbf{A}}$. Therefore, if z_p is a pole of $\mathbf{H}(z)$, i.e., $\det(z_p\mathbf{I} - \tilde{\mathbf{A}}) = 0$, it must be the eigenvalue of $\tilde{\mathbf{A}}$. The system $\mathbf{H}(z)$ is stable if all the eigenvalues of $\tilde{\mathbf{A}}$ have magnitudes less than one. In other words, we can determine whether the designed compensation system is stable by checking the eigenvalues of $\tilde{\mathbf{A}}$. Let's further define another matrix of dimension of $(N_b \times N_b)$, $\tilde{\mathbf{A}}'$, as

$$\tilde{\mathbf{A}}' = \mathbf{A}'_{M-1} \mathbf{A}'_{M-2} \dots \mathbf{A}'_1 \mathbf{A}'_0, \quad (4.27)$$

where

$$\mathbf{A}'_n = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -b_n[N_b] & -b_n[N_b - 1] & \dots & -b_n[1] \end{bmatrix}. \quad (4.28)$$

Denote the set of eigenvalues of $\tilde{\mathbf{A}}'$ as $\{\lambda_1, \dots, \lambda_{N_b}\}$. It can be verified that the set of eigenvalues of $\tilde{\mathbf{A}}$ is the set of $\{\lambda_1, \lambda_2, \dots, \lambda_{N_b}\}$ union with $(N_a - 1)$ zeros. The detailed derivation can be found in Appendix B. Therefore, the stability of the M -PTV IIR filter depends solely on the denominator coefficients. It suffices to check the eigenvalues of $\tilde{\mathbf{A}}'$ to determine whether a stable design has been reached.

4.2.2 MIMO Representation of an M -PTV FIR Filter

The equivalent MIMO representation is valid for any discrete-time PTV system. If an M -PTV FIR filter described by the following input-output relationship was to be written in the state space representation

$$y[n] = \sum_{p=0}^{N_a-1} a_n[p]x[n-p], \quad (4.29)$$

the state space parameters $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}_n$, and d_n will be modified to

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{(N_a-2) \times 1} & \mathbf{I}_{(N_a-2)} \\ 0 & \mathbf{0}_{1 \times (N_a-2)} \end{bmatrix}, \quad (4.30)$$

$$\mathbf{b} = [\mathbf{0}_{1 \times (N_a-2)}, 1]^T, \quad (4.31)$$

$$\mathbf{c}_n = [a_n[N_a-1], a_n[N_a-2], \dots, a_n[1]]^T, \quad (4.32)$$

$$d_n = a_n[0], \quad (4.33)$$

where $\mathbf{0}_{m \times n}$ denotes the $m \times n$ zero matrix/vector and $\mathbf{I}_m$ denotes the $m \times m$ identity matrix. $\mathbf{A}$ is an $(N_a - 1) \times (N_a - 1)$ matrix, $\mathbf{b}$ and $\mathbf{c}_n$ are column vectors with dimension $(N_a - 1) \times 1$ and d_n is a scalar. Note that only $\mathbf{c}_n$ and d_n are time-varying in the case of a time-varying FIR filter. Following the same derivation steps in Appendix A, the equivalent LTI MIMO system of the M -PTV FIR filter is obtained and the state space parameters of the MIMO system are given by

$$\tilde{\mathbf{A}} = \mathbf{A}^M, \quad (4.34)$$

$$\tilde{\mathbf{B}} = [\mathbf{A}^{M-1}\mathbf{b}, \dots, \mathbf{A}\mathbf{b}, \mathbf{b}], \quad (4.35)$$

$$\tilde{\mathbf{C}} = [\mathbf{c}_0, \mathbf{A}^T \mathbf{c}_1, \dots, [\mathbf{A}^{M-1}]^T \mathbf{c}_{M-1}]^T, \quad (4.36)$$

and $\tilde{\mathbf{D}}$ is a lower triangular matrix with its element at i -th row and j -th column $\tilde{d}_{ij}$ given by, assuming the indices i and j starts from zero

$$\tilde{d}_{ij} = \begin{cases} 0 & i < j \\ d_i & i = j \\ \mathbf{c}_i^T \mathbf{b} & i = j + 1 \\ \mathbf{c}_i^T \mathbf{A}^{i-j-1} \mathbf{b} & i \geq j + 2 \end{cases} \quad (4.37)$$

The state transition matrix $\tilde{\mathbf{A}}$ of the equivalent MIMO representation of an M -PTV FIR filter has all zero eigenvalues. A PTV FIR filter is always stable.

4.3 Impulse Response of an M -PTV IIR Filter

The impulse response of an M -PTV IIR filter depends on when the impulse signal is applied. The impulse response can be determined with the help of the equivalent MIMO representation [134]. Substituting (4.12) to (4.16) repeatedly, the output vector of the MIMO system can be expressed in terms of the input vector and the initial state vector as

$$\mathbf{y}[k] = \begin{cases} \tilde{\mathbf{C}}\mathbf{t}[0] + \tilde{\mathbf{D}}\mathbf{x}[0], & k = 0 \\ \tilde{\mathbf{C}}\tilde{\mathbf{A}}^k\mathbf{t}[0] + \sum_{i=1}^k \tilde{\mathbf{C}}\tilde{\mathbf{A}}^{i-1}\tilde{\mathbf{B}}\mathbf{x}[k-i] + \tilde{\mathbf{D}}\mathbf{x}[k], & k > 0 \end{cases} \quad (4.38)$$

Assuming the M -PTV IIR filter is excited by an impulse signal applied at the i -th clock cycle where $0 \leq i \leq M-1$. This is equivalent to applying the following signal vector to the input of the equivalent MIMO system

$$\mathbf{x}[k] = \begin{cases} \mathbf{e}_i, & k = 0 \\ \mathbf{0}_{M \times 1}, & k > 0 \end{cases} \quad (4.39)$$

where $\mathbf{e}_i$ is an $M \times 1$ column vector with its i -th entry being one and the rest being zero. With the initial state vector $\mathbf{t}[0] = \mathbf{0}_{N_s \times 1}$, the output vector of the MIMO system due to the impulse signal at the i -th input port, $\mathbf{r}_i[k]$, is then

$$\mathbf{r}_i[k] = \begin{cases} \tilde{\mathbf{D}}\mathbf{e}_i, & k = 0 \\ \tilde{\mathbf{C}}\tilde{\mathbf{A}}^{k-1}\tilde{\mathbf{B}}\mathbf{e}_i, & k > 0 \end{cases} \quad (4.40)$$

Consider the eigenvalue decomposition of $\tilde{\mathbf{A}}$ in the following equation

$$\tilde{\mathbf{A}} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}, \quad (4.41)$$

where $\mathbf{Q}$ is a matrix with its columns composed of the eigenvectors of $\tilde{\mathbf{A}}$ and $\mathbf{\Lambda}$ is an $N_s \times N_s$ diagonal matrix given by

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{N_s} \end{bmatrix}, \quad (4.42)$$

where $\lambda_1, \lambda_2, \dots, \lambda_{N_s}$ are the eigenvalues of $\tilde{\mathbf{A}}$ and $\lambda_p = 0$ for $N_b + 1 \leq p \leq N_s$. Substituting (4.41)

into (4.40) and letting $\mathbf{C}' = \tilde{\mathbf{C}}\mathbf{Q}$, $\mathbf{B}' = \mathbf{Q}^{-1}\tilde{\mathbf{B}}$, (4.40) becomes

$$\mathbf{r}_i[k] = \begin{cases} \tilde{\mathbf{D}}\mathbf{e}_i, & k = 0 \\ \mathbf{C}'\mathbf{\Lambda}^{k-1}\mathbf{B}'\mathbf{e}_i, & k > 0 \end{cases} \quad (4.43)$$

Let $r_{m,i}[k]$ denote the signal at the m -th output port due to the impulse signal at the i -th input port and

$r_{m,i}[k]$ can be calculated by using the following equation

$$r_{m,i}[k] = \begin{cases} \tilde{d}_{mi}, & k = 0 \\ \sum_{p=1}^{N_b} \lambda_p^{k-1} c'_{m,(p-1)} b'_{(p-1),i}, & k > 0 \end{cases} \quad (4.44)$$

where b'_{ij} , c'_{ij} and $\tilde{d}_{ij}$ are the elements at the i -th row and j -th column of the matrices $\mathbf{B}'$, $\mathbf{C}'$, and $\tilde{\mathbf{D}}$, respectively and the indices i and j start from zero. Take the z-transform of $r_{m,i}[k]$ and we have

$$\begin{aligned} R_{m,i}(z) &= \sum_{k=0}^{\infty} r_{m,i}^i[k] z^{-k} \\ &= \tilde{d}_{mi} + \sum_{k=1}^{\infty} \sum_{p=1}^{N_b} \lambda_p^{k-1} c'_{m,(p-1)} b'_{(p-1),i} z^{-k} \\ &= \tilde{d}_{mi} + \sum_{p=1}^{N_b} \frac{c'_{m,(p-1)} b'_{(p-1),i} z^{-1} (1 - (\lambda_p z^{-1})^{\infty})}{1 - \lambda_p z^{-1}}. \end{aligned} \quad (4.45)$$

Note that the MIMO system is operating at one- M th of the rate of the M -PTV IIR filter. The impulse response of the M -PTV filter due to the impulse signal at the i -th clock cycle is formed by multiplexing the M sequences, $r_{m,i}[k]$, $m = 0, 1, \dots, M-1$. Denote the z-transform of the output of the M -PTV filter due to the impulse signal at the i -th clock cycle as $R_i(z)$. $R_i(z)$ is formed by first

interpolating $R_{m,i}(z)$, $m = 0, 1, \dots, M-1$ to z^M , then delaying it by mT_s , and finally summing the M channels together, i.e.,

$$\begin{aligned} R_i(z) &= \sum_{m=0}^{M-1} R_{m,i}(z^M) z^{-m} \\ &= \sum_{m=0}^{M-1} \left(\tilde{a}_{mi} + \sum_{p=1}^{N_b} \frac{c'_{m,(p-1)} b'_{(p-1),i} z^{-M} (1 - (\lambda_p z^{-M})^\infty)}{1 - \lambda_p z^{-M}} \right) z^{-m}. \end{aligned} \quad (4.46)$$

If the designed IIR compensation filter is stable meaning that $|\lambda_p| < 1$ for $1 \leq p \leq N_b$, (4.46) is reduced to

$$\begin{aligned} R_i(z) &= \sum_{m=0}^{M-1} \left(\tilde{a}_{mi} + \sum_{p=1}^{N_b} \frac{c'_{m,(p-1)} b'_{(p-1),i} z^{-M}}{1 - \lambda_p z^{-M}} \right) z^{-m} \\ &= \frac{N_i(z)}{D(z^M)}, \end{aligned} \quad (4.47)$$

where

$$D(z^M) = \prod_{p=1}^{N_b} (z^M - \lambda_p), \quad (4.48)$$

and $N_i(z)$ is a polynomial in z . The proposed M -PTV IIR filter can equivalently be realized as a filter bank consisting of M LTI IIR filters sharing the common denominator $D(z^M)$. Alternatively, $R_i(z)$ can be expressed in terms of the elements of the transfer function matrix $\mathbf{H}(z)$ as

$$R_i(z) = \sum_{m=0}^{M-1} H_{m,i}(z^M) z^{-m}. \quad (4.49)$$

Figure 4.3 shows the alternative realization of the proposed compensation structure using the M LTI IIR filters, $R_m(z)$, $m = 0, 1, \dots, M-1$. It is readily observed that the denominator $D(z^M)$ is the characteristic polynomial of the matrix $\tilde{\mathbf{A}}'$ defined in (4.27). The poles of $R_m(z)$, $m = 0, 1, \dots, M-1$

are located at $z = \lambda_p^{1/M} e^{j2\pi k/M}$ where $p = 1, 2, \dots, N_b$ and $k = 0, 1, \dots, M-1$. The stability of $1/D(z^M)$ and hence the stability of the M -PTV IIR filter is ensured when all the eigenvalues of $\tilde{\mathbf{A}}'$ have magnitudes less than one, i.e., $|\lambda_p| < 1$ for $1 \leq p \leq N_b$.

Figure 4.3 Equivalent filter bank realization of the M -PTV IIR filter.

4.4 Alternative Approach to Derive the MIMO Transfer Function Matrix

In this section, an alternative approach to derive the transfer function matrix without using the block form state space representation is presented. The TIADC itself is an M -PTV system where any two output samples are contributed by the same sub-ADC if they are spaced by integer multiples of M samples apart and different sub-ADCs otherwise. We add a superscript m to the high rate sequence $x[n]$ to indicate that the n -th sample of $x[n]$ is contributed by Channel- m , i.e. $x^m[n]$, where $m = \langle n \rangle M$. $x[n]$, $x^m[n]$ and the sub-sequence of Channel- m $x_m[k]$ are related by

$$x[n] = x^{\langle n \rangle M}[n] = x_{\langle n \rangle M} \left[\left\lfloor \frac{n}{M} \right\rfloor \right]. \quad (4.50)$$

Similarly, denote $y^m[n]$ as the n -th sample of $y[n]$ contributed by Channel- m . $x^m[n]$ and $y^m[n]$ are clocked at the high data rate of f_s samples per second. Recall that the time-domain relationship of the

time-varying IIR compensation filter is

$$y[n] = \sum_{p=0}^{N_a-1} a_n[p]x[n-p] - \sum_{q=1}^{N_b} b_n[q]y[n-q]. \quad (4.51)$$

With the two new notations, $x^m[n]$ and $y^m[n]$, (4.51) can be rewritten as

$$y^m[n] = \sum_{p=0}^{N_a-1} a_m[p]x^{\langle m-p \rangle M}[n-p] - \sum_{q=1}^{N_b} b_m[q]y^{\langle m-q \rangle M}[n-q]. \quad (4.52)$$

For an M -channel system, there are M such input-output equations. Rearrange each equation such that the input data contributed by the same channel are grouped together and the output data contributed by the same channel are grouped together. We have

$$\begin{aligned} & \sum_{k=0}^{M-1} \sum_{q=0}^{N_{b,k}} b_0[qM+k]y^{\langle -k \rangle M}[n-(qM+k)] \\ &= \sum_{k=0}^{M-1} \sum_{p=0}^{N_{a,k}} a_0[pM+k]x^{\langle -k \rangle M}[n-(pM+k)], \\ & \vdots \\ & \sum_{k=0}^{M-1} \sum_{q=0}^{N_{b,k}} b_m[qM+k]y^{\langle m-k \rangle M}[n-(qM+k)] \\ &= \sum_{k=0}^{M-1} \sum_{p=0}^{N_{a,k}} a_m[pM+k]x^{\langle m-k \rangle M}[n-(pM+k)], \\ & \vdots \\ & \sum_{k=0}^{M-1} \sum_{q=0}^{N_{b,k}} b_{M-1}[qM+k]y^{\langle M-1-k \rangle M}[n-(qM+k)] \\ &= \sum_{k=0}^{M-1} \sum_{p=0}^{N_{a,k}} a_{M-1}[pM+k]x^{\langle M-1-k \rangle M}[n-(pM+k)]. \end{aligned} \quad (4.53)$$

In (4.53), $b_m[0] = 1, m = 0, 1, \dots, M-1$, $N_{b,k} = \lceil (N_b + 1 - k)/M \rceil$ and $N_{a,k} = \lceil (N_a - k)/M \rceil$ where $\lceil a \rceil$ is the minimum integer greater than or equal to a . Taking z-transform of (4.53) and we have

$$\begin{aligned}
& \sum_{k=0}^{M-1} z^{-k} Y^{(-k)M}(z) \sum_{q=0}^{N_{b,k}} b_0[qM + k] z^{-qM} = \sum_{k=0}^{M-1} z^{-k} X^{(-k)M}(z) \sum_{p=0}^{N_{a,k}} a_0[pM + k] z^{-pM}, \\
& \vdots \\
& \sum_{k=0}^{M-1} z^{-k} Y^{(m-k)M}(z) \sum_{q=0}^{N_{b,k}} b_m[qM + k] z^{-qM} = \sum_{k=0}^{M-1} z^{-k} X^{(m-k)M}(z) \sum_{p=0}^{N_{a,k}} a_m[pM + k] z^{-pM}, \\
& \vdots \\
& \sum_{k=0}^{M-1} z^{-k} Y^{(M-1-k)M}(z) \sum_{q=0}^{N_{b,k}} b_{M-1}[qM + k] z^{-qM} \\
& \quad = \sum_{k=0}^{M-1} z^{-k} X^{(M-1-k)M}(z) \sum_{p=0}^{N_{a,k}} a_{M-1}[pM + k] z^{-pM}.
\end{aligned} \tag{4.54}$$

Rewrite (4.54) in matrix-vector form

$$\begin{bmatrix} Y^0(z) \\ Y^1(z) \\ \vdots \\ Y^2(z) \end{bmatrix} = \mathbf{G}(z) \begin{bmatrix} X^0(z) \\ X^1(z) \\ \vdots \\ X^2(z) \end{bmatrix}, \tag{4.55}$$

where

$$\mathbf{G}(z) = \mathbf{B}^{-1}(z) \mathbf{A}(z), \tag{4.56}$$

and $\mathbf{B}(z)$ and $\mathbf{A}(z)$ are $M \times M$ polynomial matrices in the form of

$$\mathbf{B}(z) = \begin{bmatrix} B_{00}(z) & B_{01}(z) & \cdots & B_{0,M-1}(z) \\ B_{10}(z) & B_{11}(z) & \cdots & B_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ B_{M-1,0}(z) & B_{M-1,1}(z) & \cdots & B_{M-1,M-1}(z) \end{bmatrix}, \tag{4.57}$$

$$\mathbf{A}(z) = \begin{bmatrix} A_{00}(z) & A_{01}(z) & \cdots & A_{0,M-1}(z) \\ A_{10}(z) & A_{11}(z) & \cdots & A_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ A_{M-1,0}(z) & A_{M-1,1}(z) & \cdots & A_{M-1,M-1}(z) \end{bmatrix}. \quad (4.58)$$

The elements at the i -th row and the j -th column of $\mathbf{B}(z)$ and $\mathbf{A}(z)$ are given by, assuming the indices i and j start from zero

$$B_{ij}(z) = z^{-k} \sum_{q=0}^{N_{b,k}} b_i[qM + k]z^{-qM}, k = \langle i - j \rangle M, \quad (4.59)$$

$$A_{ij}(z) = z^{-k} \sum_{p=0}^{N_{a,k}} a_i[pM + k]z^{-pM}, k = \langle i - j \rangle M. \quad (4.60)$$

Note that $\mathbf{B}(z)$ is reduced to an identity matrix in the FIR case. Comparing to the transfer function matrix of the equivalent MIMO system, $\mathbf{H}(z)$, derived in Section 4.2.1, $\mathbf{G}(z)$ is operating at the high data rate f_s whereas $\mathbf{H}(z)$ is operating at f_s/M . The signal at the m -th input port of $\mathbf{G}(z)$ is the subsequence of ADC- m being up-sampled by a factor of M and delayed by m samples. The equivalent high rate MIMO system of the M -PTV filter is shown in Figure 4.4.

Figure 4.4 Equivalent high rate MIMO system of the M -PTV IIR filter.

Further write $\mathbf{G}(z)$ as

$$\mathbf{G}(z) = \frac{\text{adj}(\mathbf{B}(z)) \cdot \mathbf{A}(z)}{\det(\mathbf{B}(z))}, \quad (4.61)$$

where $\text{adj}(\mathbf{B}(z))$ and $\det(\mathbf{B}(z))$ denote the adjugate and the determinant of $\mathbf{B}(z)$, respectively. The elements of $\mathbf{G}(z)$ share the common denominator $\det(\mathbf{B}(z))$. Let $G_{i,j}(z)$ be the entry of $\mathbf{G}(z)$ at the i -th row and the j -th column. $G_{i,j}(z)$ represents the z -transform of the i -th output due to an impulse signal at the j -th input port. Assume the M -PTV IIR filter is excited by an impulse signal applied at the i -th clock cycle where $0 \leq i \leq M - 1$. This is equivalent to exciting the i -th input port of $\mathbf{G}(z)$ at the i -th clock cycle. The z -transform of the signal at the m -th output port is then $z^{-i}G_{m,i}(z)$. This signal is also the response at the m -th output port of $\mathbf{H}(z)$ being up-sampled by a factor of M and delayed by m samples. Thus the elements of $\mathbf{G}(z)$ and $\mathbf{H}(z)$ are related according to

$$z^{-i}G_{m,i}(z) = z^{-m}H_{m,i}(z^M). \quad (4.62)$$

It is clear that the poles of $\mathbf{G}(z)$ are the poles of $\mathbf{H}(z^M)$ and are located at $z = \lambda_p^{1/M} e^{j2\pi k/M}$, $p = 1, \dots, N_s$, $k = 0, 1, \dots, M - 1$. The stability criterion of the proposed IIR compensation filter can be derived by calculating the determinant of the matrix $\mathbf{B}(z)$ given in (4.57). The z -transform of the output of the M -PTV filter due to the impulse signal at the i -th clock cycle, $R_i(z)$, can be expressed in terms of the elements of $\mathbf{G}(z)$ as

$$R_i(z) = z^{-i} \sum_{m=0}^{M-1} G_{m,i}(z). \quad (4.63)$$

(4.63) is another way to express the transfer function of the LTI IIR filters in the equivalent filter bank structure shown in Figure 4.3.

Example 4.1

We shall use an example of an M -PTV IIR filter where $M = 3$ to demonstrate the relationship between the two transfer function matrices $\mathbf{H}(z)$ and $\mathbf{G}(z)$. Assume the filter coefficients of the PTV IIR filter are those listed in Table 4.1.

Table 4.1 Filter coefficients of the PTV IIR filter in Example 4.1.

	$m = 0$	$m = 1$	$m = 2$
$a_m[0]$	1.05	0.90	0.85
$a_m[1]$	0.04	0.03	0.06
$a_m[2]$	0.02	0.05	0.01
$b_m[1]$	0.10	0.15	0.05
$b_m[2]$	0.20	0.25	0.40

Following the procedures in Section 4.2.1 and this section, the transfer function matrices $\mathbf{H}(z)$ and $\mathbf{G}(z)$ are derived from the filter coefficients and they are

$$\mathbf{H}(z)$$

$$= \frac{1}{1 - 0.08325z^{-1} - 0.02z^{-2}} \cdot \begin{bmatrix} 1.05 - 0.005975z^{-1} + 0.0005z^{-2} & 0.2305z^{-1} + 0.0025z^{-2} & 0.12z^{-1} + 0.04975z^{-2} \\ 0.1875 + 0.10645z^{-1} & 0.9 - 0.0141z^{-1} + 0.004z^{-2} & 0.268z^{-1} + 0.002z^{-2} \\ 0.439375 + 0.0021z^{-1} & 0.105 + 0.0865z^{-1} & 0.8 - 0.0052z^{-1} + 0.004z^{-2} \end{bmatrix}$$

$$\mathbf{G}(z)$$

$$= \frac{1}{1 - 0.08325z^{-3} - 0.02z^{-6}} \cdot \begin{bmatrix} 1.05 - 0.005975z^{-3} + 0.0005z^{-6} & 0.2305z^{-2} + 0.0025z^{-5} & 0.12z^{-1} + 0.04975z^{-4} \\ 0.1875z^{-1} + 0.10645z^{-4} & 0.9 - 0.0141z^{-3} + 0.004z^{-6} & 0.268z^{-2} + 0.002z^{-5} \\ 0.439375z^{-2} + 0.0021z^{-5} & 0.105z^{-1} + 0.0865z^{-4} & 0.8 - 0.0052z^{-3} + 0.004z^{-6} \end{bmatrix}$$

It is clear that the elements of $\mathbf{G}(z)$ share a common denominator. The denominator of $\mathbf{G}(z)$ is the denominator of $\mathbf{H}(z)$ being up-sampled by three. The entries of $\mathbf{H}(z)$ and $\mathbf{G}(z)$ are related according to (4.62).

4.5 Chapter Summary

In this chapter, we have demonstrated that the proposed M -PTV IIR compensation filter operating at the rate of f_s can be interpreted as an equivalent LTI MIMO system operating at the rate of f_s/M . The proposed IIR compensation filter is stable if all the eigenvalues of the matrix, $\tilde{\mathbf{A}}'$, have magnitudes smaller than one. It was further shown that the M -PTV IIR filter has an equivalent IIR filter bank

representation where all the IIR filters share a common denominator. An alternative approach without using the state space representation to derive the stability criterion was also presented. Although the stability criterion can be explicitly derived, it is difficult to incorporate the eigenvalue based stability constraints into the WLS optimization process. A possible way to achieve a stable M -PTV IIR filter design is by constraining the coefficients of the common denominator of the equivalent filter bank. Since the filter bank is LTI, many existing methods for ensuring the stability of the LTI IIR filters can be applied [128-130]. However, the denominator coefficients of the equivalent filter bank are very complicated polynomials of $b_m[q]$, $m = 0, 1, \dots, M-1$, $q = 1, 2, \dots, N_b$ and the design complexity would be increased tremendously if we constrain the common denominator coefficients. Therefore, we need to search for other simpler methods for achieving a stable M -PTV IIR compensation filter design.

CHAPTER 5 : METHODS TO ENSURE THE STABILITY OF THE M -PTV IIR COMPENSATION FILTER

5.1 Introduction

The stability criterion of the M -PTV IIR filter has been derived explicitly in Chapter 4. The M -PTV IIR filter is stable if all the eigenvalues of the matrix $\tilde{\mathbf{A}}'$ given in (4.27) have magnitudes smaller than one. However, it is difficult to incorporate the eigenvalue based stability constraint into the WLS optimization process and thus the stability criterion is merely used to check whether a stable design has been reached in the design process.

For the case of $N_b = 1$, the matrix $\tilde{\mathbf{A}}'$ is simply the product of the M denominator coefficients of all the channels, i.e., $\tilde{\mathbf{A}}' = \prod_{m=0}^{M-1} b_m[1]$. When the IIR compensation filters are designed to correct for the relative mismatches in ADC- m , $m = 1, 2, \dots, M - 1$ with respect to that in ADC-0, the denominator coefficients for estimating ADC-0 output are all zeros. In this case, the matrix $\tilde{\mathbf{A}}'$ is zero and hence its eigenvalue is zero. The proposed IIR filter is always stable when the filters are designed to compensate for the relative mismatch and the number of denominator coefficient equals to one. For all the other cases, we shall introduce two simple “computer-aided” approaches for gradually achieving a stable IIR filter design. In this chapter and the subsequent chapters, we consider only the case of designing IIR filters to compensate for the relative mismatches. However, the discussion is also valid for the case of designing IIR filters to compensate for the absolute mismatch errors.

5.2 Method-1: By Increasing the Integer Group Delay

Recall that in Chapter 3, a suitable integer delay, η , has been added to the estimation equation to make the reconstruction system causal. In other words, we aim to estimate the output which is η samples delayed rather than the immediate output. ηT_s is the time difference between the “most recently available sample” from the TIADC output and the “currently computed sample”. The integer delay η represents the group delay introduced by our IIR filter reconstruction system. The choice of the group delay η affects the stability of the M -PTV IIR compensation filter. Using *Rouché’s Theorem* [135], we are able to prove that it is always possible to arrive at a stable solution by increasing the value of η .

Rouché’s Theorem: Suppose that f and g are meromorphic functions defined in the simply connected domain D , that C is a simply closed contour in D , and that f and g have no zeros or poles for $z \in C$. If the strictly inequality $|f(z) + g(z)| < |f(z)| + |g(z)|$ holds for all $z \in C$, then

$$N_{z,f} - N_{p,f} = N_{z,g} - N_{p,g}$$

where $N_{z,f}$ and $N_{p,f}$ are the number of zeros and poles respectively of f in C , and $N_{z,g}$ and $N_{p,g}$ are the number of zeros and poles respectively of g in C . All poles and zeros are counted with multiplicity.

Assume an IIR filter is designed to approximate the frequency response of the ideal filter $H_{ideal}(z)$ and the obtained filter is unstable. Denote the unstable filter as $H_{us}(z)$, then we have

$$H_{us}(e^{j\omega}) \approx H_{ideal}(e^{j\omega}). \quad (5.1)$$

Next the group delay is increased by d samples and the IIR filter is redesigned. Denote the new filter as $H_d(z)$ and we now have

$$H_d(e^{j\omega}) \approx e^{-jd\omega} H_{ideal}(e^{j\omega}), \quad (5.2)$$

or

$$e^{jd\omega}H_d(e^{j\omega}) \approx H_{ideal}(e^{j\omega}). \quad (5.3)$$

The filters designed for TIADC mismatch compensation are all-pass filters. Hence, for well approximated filters, the magnitude of the difference between $H_{us}(e^{j\omega})$ and $e^{jd\omega}H_d(e^{j\omega})$ is very small and so is smaller than the sum of the magnitudes of $H_{us}(e^{j\omega})$ and $e^{jd\omega}H_d$. We may write $|H_{us}(e^{j\omega}) - e^{jd\omega}H_d(e^{j\omega})| < |H_{us}(e^{j\omega})| + |e^{jd\omega}H_d(e^{j\omega})|$. Hence the following is true along the unit circle, i.e., at $z = e^{j\omega}$

$$|H_{us}(z) - z^d H_d(z)| < |H_{us}(z)| + |-z^d H_d(z)|. \quad (5.4)$$

The number of zeros and poles of $H_{us}(z)$ and $H_d(z)$ inside the unit circle are then related by

$$N_{z,us} - N_{p,us} = d + N_{z,d} - N_{p,d}, \quad (5.5)$$

where $N_{z,us}$ and $N_{p,us}$ are the number of zeros and poles respectively of $H_{us}(z)$ inside the unit circle, and $N_{z,d}$ and $N_{p,d}$ are the number of zeros and poles respectively of $H_d(z)$ inside the unit circle. Since $(N_{z,us} - N_{p,us})$ is fixed, increasing the value of d would force $(N_{z,d} - N_{p,d})$ to decrease. When increasing the value of d , $N_{z,d}$ may decrease. However, $N_{z,d}$ can only decrease to an extent that $N_{z,d} = 0$. Further increasing the value of d would still increase $N_{p,d}$. Therefore, it is always possible to find a big enough d such that $N_{p,d}$ reaches its maximum, which is the total number of poles of $H_d(z)$. This means that all the poles of $H_d(z)$ have been brought inside the unit circle, resulting in a stable filter design.

In the foregoing proof, the filters $H_{us}(z)$ and $H_d(z)$ are LTI IIR filters. Nonetheless, the same conclusion can be drawn for our M -PTV IIR filter. Recall that an M -PTV IIR filter has an equivalent filter bank representation. The filter bank consists of M LTI IIR filters sharing a common denominator and the poles of the IIR filter bank are the poles of the M -PTV IIR filter. If the group delay is

increased by d and the M -PTV filter is redesigned, the group delays of the M LTI IIR filters in equivalent filter bank are also increased by d . Eventually, all the poles of the common denominator can be brought inside the unit circle by increasing the group delay η and a stable solution can be obtained.

Next, we shall investigate how the stability of the IIR mismatch compensation filter is affected when the group delay η is varied from 0 to $N_a - 1$. To this end, we denote the maximum pole radius of the M -PTV IIR compensation filter as p_{max} . p_{max} is also the maximum pole radius of the equivalent IIR filter bank and the largest eigenvalue magnitude of $\tilde{\mathbf{A}}'$ to the power of $1/M$. In the design of the M -PTV FIR compensation filters, for a fixed FIR filter length, the selection of the group delay η affects the compensation performance. In general, the compensation performance is maximized when η is selected to be about half of the FIR filter length N , i.e., $\eta = (N - 1)/2$ for odd N and $\eta = N/2$ for even N . We shall also investigate how η affects the compensation performance of our M -PTV IIR filters. The IIR compensation filter may be unstable when the value of η is varied from 0 to $N_a - 1$ and the compensation performance may not be available for some values of η . Hence the WLS error rather than the actual compensation performance is studied. Recall that in Chapter 3, the compensation performance of the IIR filter is determined by not only the WLS error but also the feedback coefficients. Nevertheless, the WLS error can still be considered as a good performance indicator for the IIR compensation filter.

Example 5.1

We shall use an example to demonstrate the effect of η to the WLS error and the stability. Assume a 4-channel TIADC with timing and bandwidth mismatches shown in Table 5.1. ADC-0 is chosen as the reference and so it has no timing mismatch and its bandwidth is assumed to be Ω_s , where $\Omega_s = 2\pi f_s$ and f_s is the overall sampling rate of the TIADC. The number of feed-forward coefficients in each channel, N_a , is fixed to be 15. Five different values of N_b are simulated and the values are $N_b = 2, 3, 4, 5, 7$. For each value of N_b , the value of the group delay η is varied from 0 to $N_a - 1$. For each

set of values of N_a , N_b and η , the IIR compensation filter is optimized in the least squares sense over the frequency range of $[0, 0.45f_s]$. $w_m(\Omega)$ is assigned the value one for $0 \leq \Omega \leq 0.9\pi f_s$ and zero elsewhere.

Table 5.1 Mismatch parameters for the 4-Channel TIADC in Example 5.1

$G_0(j\Omega)$ $= \frac{1}{1 + j\Omega/\Omega_s}$	$G_1(j\Omega)$ $= \frac{e^{-j0.037\Omega T_s}}{1 + j\Omega/(1.05\Omega_s)}$	$G_2(j\Omega)$ $= \frac{e^{j0.029\Omega T_s}}{1 + j\Omega/(0.98\Omega_s)}$	$G_3(j\Omega)$ $= \frac{e^{j0.054\Omega T_s}}{1 + j\Omega/(1.03\Omega_s)}$
--	--	---	---

Define the root-mean-square (RMS) value of the WLS error as

$$e_{rms} = \sqrt{\frac{1}{LM} \sum_{m=0}^{M-1} \sum_{i=1}^L |\varepsilon_m(j\Omega_i)|^2}. \quad (5.6)$$

The values of e_{rms} in dB and the values of p_{max} for each simulation case are recorded and plotted with respect to η in Figure 5.1 and Figure 5.2, respectively.

Figure 5.1 Effect of the group delay η on the WLS error.

Figure 5.2 (a) Effect of the group delay η on the maximum pole radius p_{max} . (b) Zoomed in view where the range of η is from 6 to 14.

For each case of N_b , as the value of η is increased from 0 to $N_a - 1$, the value of e_{rms} decreases first and then rises up. For the cases where $N_b = 2$ and $N_b = 3$, e_{rms} reaches minimum when $\eta = 7$. For all the other cases of N_b , e_{rms} is minimum when $\eta = 6$. The value of η at which the minimum WLS error occurs may lead to an unstable filter design. Take the case of $N_b = 3$ for example, the smallest

value of e_{rms} is -89.1 dB occurring at $\eta = 7$. The value of p_{max} at $\eta = 7$ is 1.138 and the designed filter is not stable. Nevertheless, it is observed from Figure 5.2 that the maximum pole radius p_{max} in general decreases with increasing η . When η is increased to 10, p_{max} becomes 0.917 indicating a stable filter design; the value of e_{rms} is increased to -81.8 dB. Therefore, it is possible to arrive at a stable solution by increasing the group delay η at the expenses of larger e_{rms} .

We also observe that for the same value of η , as the number of feedback coefficients, N_b , is increased from 2 to 7, the value of e_{rms} decreases. This is expected because the estimated output is closer to the ideal value when more data are used in the estimation. However, achieving smaller value of e_{rms} by employing more feedback coefficients may not always be desirable. This is because the value of p_{max} increases with N_b , meaning that larger value of η is needed to make the filter stable which in turn resulting larger e_{rms} . Hence the compensation performance of the filter with more denominator coefficients can be worse than that with fewer denominator coefficients. For instance, when $N_b = 7$, the value of η for a stable filter design is 12 and e_{rms} is -84.1 dB. When $N_b = 5$, the value of η for a stable filter design is 11 and e_{rms} is -88.7 dB which is better than the case of $N_b = 7$ even two fewer feedback coefficients are used.

Through numerous simulations, we have found that for fixed values of N_a , N_b and mismatch errors, p_{max} decreases as the value of η increases; and that in most of the cases, the value of η which would result a stable design lies in the range of $(N_a - 1)/2 \leq \eta \leq N_a - 1$. For some cases where N_b is small, $N_b = 1$ or 2, the value of η which would result a stable design can be smaller than $(N_a - 1)/2$. We do not know the minimum value of η which would result in a stable IIR filter design. And therefore, in the design of the IIR compensation filter, a few iterations may be needed to find the suitable value of η .

5.3 Method-2: By Constraining the Magnitudes of the Denominator Coefficients

If the choice of the value for η is flexible, it is always possible to choose a value of η that will produce a stable solution. If a specific value has been pre-assigned to η and if the pre-assigned value of η does not produce a stable filter, the filter may be constrained to be stable by constraining the magnitudes of the feedback coefficients $b_m[q]$, where $q = 1, 2, \dots, N_b, m = 0, 1, \dots, M - 1$. In our approach, the magnitudes of the feedback coefficients are constrained by adding positively weighted versions of $|b_m[q]|^2$ into the least squares objective function. The larger is the weight of $|b_m[q]|^2$, the smaller will be the magnitude of $b_m[q]$. As the weight approaches infinity, the magnitude of $b_m[q]$ approaches zero and the filter becomes an FIR filter. Thus, stability can always be achieved by adding sufficiently large weight to $|b_m[q]|^2$. However, adding weighted versions of $|b_m[q]|^2$ to the least squares objective function will affect the frequency response of the filter; the larger is the weight, the more is the frequency response being affected. We do not know the minimum weight required for achieving stability and so our approach is a “computer-aided” design approach of progressively increasing the weights until stability is reached.

The procedure for adding positively weighted version of $|b_m[q]|^2$ to the least squares objective function are as follows. Let the weight of $|b_m[q]|^2$ be $w_{m,q}$. An $N_b \times N_b$ diagonal matrix $\mathbf{W}_{b,m}$ consisting of $w_{m,q}$ as the diagonal elements arranged in the same order as $b_m[q]$ in $\mathbf{h}_m$ is constructed,

$$\mathbf{W}_{b,m} = \begin{bmatrix} w_{m,1} & \mathbf{0} & 0 \\ \mathbf{0} & \ddots & \mathbf{0} \\ 0 & \mathbf{0} & w_{m,N_b} \end{bmatrix}. \quad (5.7)$$

Let $\mathbf{e}_{b,m}$ be an $N_b \times 1$ column vector containing the elements of $\sqrt{w_{m,q}} b_m[q]$ arranged in the same order as $b_m[q]$ in $\mathbf{h}_m$. Let “ $\Leftarrow$ ” denote “left side quantity” is replaced by “right-side quantity”. Using these notations, (3.25) through (3.28) can now be rewritten as

$$\boldsymbol{\varepsilon}_m \Leftarrow [\boldsymbol{\varepsilon}_m^T \quad \mathbf{e}_{b,m}^T]^T, \quad (5.8)$$

$$\mathbf{u}_m \Leftarrow [\mathbf{u}_m^T \quad \mathbf{0}_{1 \times N_b}]^T, \quad (5.9)$$

$$\mathbf{s}_m \Leftarrow \begin{bmatrix} \mathbf{S}_m \\ [\mathbf{0}_{N_b \times N_a} \quad \mathbf{I}_{N_b}] \end{bmatrix}, \quad (5.10)$$

$$\mathbf{W}_m \Leftarrow \begin{bmatrix} \mathbf{W}_m & \mathbf{0}_{L \times N_b} \\ \mathbf{0}_{N_b \times L} & \mathbf{W}_{b,m} \end{bmatrix}, \quad (5.11)$$

where $\mathbf{0}_{i \times j}$ denotes the i -row by j -column null matrix/vector and $\mathbf{I}_i$ denotes the identity matrix of size $i \times i$. (5.8) through (5.11) are used instead of (3.25) through (3.28) in (3.30) to solve for $\mathbf{h}_m$.

Example 5.2

We shall use an example to show that the stability of the IIR filter can be achieved by constraining the magnitudes of the denominator coefficients. Consider the 4-channel TIADC whose mismatch parameters are listed in Table 5.2. The number of numerator coefficients in each channel, N_a , is fixed to 19 and value of the group delay η is fixed to 10. Three different values of N_b are simulated and the values are $N_b = 2, 5, 8$. To facilitate comparison, all the denominator coefficients in each design case are assigned the same weight. The uniform weight is varied from 0 to 2×10^{-5} with a step size of 2×10^{-6} . The filters are optimized in the least squares sense over the frequency range of $[0, 0.45f_s]$.

Table 5.2 Mismatch parameters for the 4-channel TIADC in Example 5.2

$G_0(j\Omega)$ $= \frac{1}{1 + j\Omega/\Omega_s}$	$G_1(j\Omega)$ $= \frac{e^{j0.087\Omega T_s}}{1 + j\Omega/(1.09\Omega_s)}$	$G_2(j\Omega)$ $= \frac{e^{-j0.052\Omega T_s}}{1 + j\Omega/(0.85\Omega_s)}$	$G_3(j\Omega)$ $= \frac{e^{j0.068\Omega T_s}}{1 + j\Omega/(1.12\Omega_s)}$
--	---	--	---

The RMS WLS error e_{rms} in dB and the maximum pole radius p_{max} versus the denominator coefficient weight are plotted in Figure 5.3 and 5.4, respectively. For each simulation case, as the denominator coefficient weight increases gradually, e_{rms} increases and p_{max} decreases. When no weight is imposed on the denominator, the filter with $N_b = 2$ is stable and the maximum pole radius is 0.973 whereas the values of p_{max} are 1.296 and 1.566, respectively, for $N_b = 5$ and $N_b = 8$. With

$N_b = 5$, the filter is stable when the weight is increased to 4×10^{-6} and the value of e_{rms} is -76.0 dB. Greater weight of 8×10^{-6} is needed for the filter with $N_b = 8$ to be stable and the value of e_{rms} is -75.4 dB. Therefore, using more denominator coefficients does not always result in better compensation performance.

Figure 5.3 Effect of the denominator coefficient weight on the WLS error.

Figure 5.4 Effect of the denominator coefficient weight on the maximum pole radius p_{max} .

5.4 Minimum Group Delay Required to Achieve Stability

Denote the minimum value of η required to achieve stability as η_{min} . We shall perform a thorough investigation on the relationships between η_{min} and the various design parameters including the number of numerator and denominator coefficients, N_a and N_b , the number of interleaving channels, M , the scale of the mismatch errors, ε_{max} . Two situations shall be studied; one is the case where only timing mismatch is present and the other is the case where only bandwidth mismatch is present. We shall derive explicit formulas expressing η_{min} in terms of N_a , N_b , M and ε_{max} . The results can be used as general guidelines for selecting a suitable value of η to achieve a stable IIR compensation filter design.

5.4.1 Timing Mismatch Case

Assume the timing mismatch in Channel- m , $m = 1, 2, \dots, M - 1$ is within the range of $(-\varepsilon_{max}, \varepsilon_{max})$ where ε_{max} indicates the scale of the timing mismatch error. ADC-0 is chosen as the reference and thus assumed to have no timing mismatch. In the experiment, the value of ε_{max} is varied from 0.01 to 0.1 with a step size of 0.01. For each value of ε_{max} , $(M - 1)$ random numbers are drawn from the uniform distribution over the interval of $(-\varepsilon_{max}, \varepsilon_{max})$. The $(M - 1)$ random numbers represent the relative timing mismatches in ADC- m , $m = 1, 2, \dots, M - 1$ with respect to ADC-0. This process is repeated 10 thousand times so that a total of 10 thousand sets of timing mismatches are generated for each value of ε_{max} . Then, for each set of the timing mismatch values, IIR filters are designed using various combinations of N_a and N_b . The value of N_a is varied from 10 to 30 and the value of N_b is varied from 2 to 10 for each value of N_a . In each design, the value of η is varied from 0 to $N_a - 1$ to identify the minimum group delay required to achieve stability, i.e., η_{min} . The numbers of interleaving channels simulated are $M = 2, 4$, and 5. All the IIR filters are optimized in the least squares over the frequency range of $[0, 0.45f_s]$. Table 5.3 summarizes all the design parameters and their values simulated in the experiment.

Table 5.3 List of the values of the design parameters simulated to identify η_{min} . Timing mismatch case.

Optimization frequency range : $[0, 0.45f_s]$	
M	: 2, 4, 5
ε_{max}	: 0.01, 0.02, ..., 0.1
N_a	: 10, 11, ..., 30
N_b	: 2, 3, ..., 10

There are a total of 5670 different combinations of M , ε_{max} , N_a , and N_b and for each combination, 10 thousand cases of η_{min} value are obtained. We found that in more than 99% of the 10 thousand cases, the value of η_{min} falls on k consecutive numbers where k is ranging from one to eight for different combinations of M , ε_{max} , N_a , and N_b . The occurrence of k in the 5670 combinations are counted and the histogram of k is plotted in Figure 5.5. It can be seen from Figure 5.5 that for a majority of the design parameter combinations simulated, the value of η_{min} falls on a single number.

Figure 5.5 Occurrence of k in the 5670 different sets of design parameters where k is the range of consecutive numbers that more than 99% of the 10 thousand cases of η_{min} fall on for each set of design parameters. Timing mismatch case.

To simplify the analysis, we take the mean value of η_{min} in the 10 thousand runs and denote the mean value as $\bar{\eta}_{min}$ and wish to express $\bar{\eta}_{min}$ as a function of M , ε_{max} , N_a , and N_b as follows

$$\bar{\eta}_{min} \approx f(M, \varepsilon_{max}, N_a, N_b). \quad (5.12)$$

From the above expression, we then wish to derive the range of numbers that η_{min} could fall on, i.e.,

$$\eta_{min} = [\text{round}(\bar{\eta}_{min}) - \Delta\eta, \text{round}(\bar{\eta}_{min}) + \Delta\eta], \quad (5.13)$$

where $\text{round}(\cdot)$ denotes the operation of rounding “ $\cdot$ ” to the nearest integer and $\Delta\eta$ measures the span of the numbers for searching for η_{min} . The rounding error as well as the error introduced by curve fitting is also absorbed by $\Delta\eta$. The number of values to search for η_{min} is narrowed down to $(2\Delta\eta + 1)$ from N_a . As will be shown later, $\Delta\eta$ equals to one in most of the design cases. That means no more than three values of η need to be tested to obtain a stable design.

The relationship between $\bar{\eta}_{min}$ and N_a is first investigated. Figure 5.6 shows the plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and ε_{max} while keeping the number of interleaving channels, M , fixed to 2. In each subplot, the value of N_b is fixed to a specific value whereas the value of ε_{max} is varied from 0.01 to 0.1. It can be seen from each subplot that the 10 curves corresponding to $\varepsilon_{max} = 0.01, 0.02, \dots, 0.1$ almost coincide with each other and therefore we conclude that the effect of ε_{max} to $\bar{\eta}_{min}$ is very weak when $M = 2$ and $\varepsilon_{max} \leq 0.1$. It is also observed that the relationship between $\bar{\eta}_{min}$ and N_a is approximately linear. In Figure 5.7, the value of ε_{max} is fixed to 0.01, and $\bar{\eta}_{min}$ is plotted against N_a for various combinations of N_b and M . $\bar{\eta}_{min}$ and N_a are approximately linearly related regardless of the number of interleaving channels. The value of $\bar{\eta}_{min}$ generally increases with increasing value of M for fixed values of N_a , N_b , and ε_{max} .

Figure 5.6 Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and ϵ_{max} , M is fixed to 2. Timing mismatch case.

Figure 5.7 Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and M , ϵ_{max} is fixed to 0.01. Timing mismatch case.

Figure 5.8 Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and ϵ_{max} , M is fixed to 4. Timing mismatch case.

Figure 5.9 Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and M , ϵ_{max} is fixed to 0.05. Timing mismatch case.

Next we consider the effect of the number of denominator coefficients N_b on $\bar{\eta}_{min}$. The value of $\bar{\eta}_{min}$ is plotted against N_b in Figure 5.8 for different combinations of N_a and ε_{max} while keeping the value of M fixed to 4. In each subplot, the value of N_a is fixed to a specific value and ε_{max} is varied from 0.01 to 0.1 with a step size of 0.01. The value of $\bar{\eta}_{min}$ increases with N_b and the plotted curves are in the shape of logarithmic functions. In each subplot, the value of $\bar{\eta}_{min}$ increases as ε_{max} increases. However, for $N_a < 26$, the 10 curves corresponding to $\varepsilon_{max} = 0.01, 0.02, \dots, 0.1$ almost coincide with each other and the effect of ε_{max} on $\bar{\eta}_{min}$ is negligible. Figure 5.9 plots $\bar{\eta}_{min}$ versus N_b for different values of M while keeping $\varepsilon_{max} = 0.05$. $\bar{\eta}_{min}$ in general increases with increasing M .

Figure 5.10 Plots of $\bar{\eta}_{min}$ versus ε_{max} for different values of N_a . Timing mismatch case. (a) The case where $N_b = 3$ and $M = 2$. (b) The case where $N_b = 9$ and $M = 5$.

In Figure 5.10, the value of $\bar{\eta}_{min}$ is plotted against ε_{max} where ε_{max} is varied from 0.01 to 0.1. Figure 5.10(a) shows the case where $N_b = 3$ and $M = 2$ and Figure 5.10(b) shows the case where $N_b = 9$ and $M = 5$. There are 21 curves in each plot corresponding to $N_a = 10, 22, \dots, 30$. In Figure 5.10(a), the value of $\bar{\eta}_{min}$ remains almost constant as ε_{max} increases from 0.01 to 0.1 for all the values of N_a . In Figure 5.10(b), the effect of ε_{max} on $\bar{\eta}_{min}$ becomes noticeable only when $N_a \geq 26$. From the

numerous simulation results, we observe that ε_{max} has a noticeable effect on $\bar{\eta}_{min}$ only when the values of N_a , N_b , and M are large, hence the effect of ε_{max} on $\bar{\eta}_{min}$ is only secondary and is insignificant compared to that of N_a , N_b , and M on $\bar{\eta}_{min}$. Therefore, to simplify the formula, we shall ignore the effect of ε_{max} in the derivation. The error caused by this simplification can be absorbed into the term $\Delta\eta$ in (5.13). The objective now becomes finding an explicit formula expressing $\bar{\eta}_{min}$ in terms of N_a , N_b , and M , i.e., $\bar{\eta}_{min} = f(M, N_a, N_b)$.

To do this, the values of N_b and M are fixed first, and $\bar{\eta}_{min}|_{N_b, M \text{ fixed}} = f(N_a)$ is identified for every combination of N_b and M . It is observed from Figure 5.6 and 5.7 that, the value of $\bar{\eta}_{min}$ increases almost linearly with N_a . Assume $\bar{\eta}_{min}$ and N_a have the following linear relationship,

$$\bar{\eta}_{min}|_{N_b, M \text{ fixed}} \approx u_0 + u_1 N_a. \quad (5.14)$$

The coefficients u_0 and u_1 are then determined using line fitting method for each combination of N_b and M . Next, for each value of M , we attempt to express the coefficients u_l where $l = 0, 1$ as functions of N_b . It is observed in Figure 5.8 and 5.9 that $\bar{\eta}_{min}$ increases roughly logarithmically with N_b and hence we assume the coefficients $u_l, l = 0, 1$ are logarithmic functions of N_b . We have compared several different logarithmic functions of N_b in approximating $u_l, l = 0, 1$ and found that the error is the smallest when the coefficients are approximated using polynomials of $\ln(\ln(N_b))$. Assume that the coefficient u_l is expressed in terms of N_b as follows

$$u_l \approx \sum_{m=0}^p u_{l,m} (\ln(\ln(N_b)))^m, \quad l = 0, 1. \quad (5.15)$$

The coefficients $u_{l,m}$ where $m = 0, 1, \dots, p$ are then determined using curve fitting method for each value of M . The approximation error decreases as higher order terms of $\ln(\ln(N_b))$ are included. When $p \geq 2$, further increasing the value of p only results a marginal decrement of the approximation error. Hence we choose $p = 2$ to keep the formula relatively simple. To summarize, the value of $\bar{\eta}_{min}$ for a specific number of interleaving channels, can be estimated using the following equation

$$\bar{\eta}_{min}|_{M \text{ fixed}} \approx u_{0,0} + u_{1,0}N_a + (u_{0,1} + u_{1,1}N_a)\ln(\ln(N_b)) + (u_{0,2} + u_{1,2}N_a)(\ln(\ln(N_b)))^2. \quad (5.16)$$

From Figure 5.7 and 5.9, it is observed that larger value of η would be needed to achieve a stable IIR filter design when the number of interleaving channels is increased. However, only three cases of M have been simulated and more data points would be needed to determine the relationship between $\bar{\eta}_{min}$ and M . Here we only list the coefficient values of $u_{l,m}$, $l = 0, 1$, $m = 0, 1, 2$ for $M = 2, 4$ and 5 in Table 5.4. Using (5.16) and the coefficients listed in Table 5.4, the RMS error in approximating $\bar{\eta}_{min}$ are 0.636, 0.354 and 0.290 for $M = 2, 4$ and 5 respectively.

Table 5.4 The coefficients in (5.16) for approximating $\bar{\eta}_{min}$. Timing mismatch case.

	$u_{0,0}$	$u_{1,0}$	$u_{0,1}$	$u_{1,1}$	$u_{0,2}$	$u_{1,2}$
$M = 2$	-1.267	0.535	0.314	0.175	0.612	0.174
$M = 4$	-0.375	0.618	0.677	0.308	-1.928	0.018
$M = 5$	-0.311	0.633	-0.570	0.356	-0.590	-0.048

The last step is to identify the values of $\Delta\eta$ such that for each combination of M , ε_{max} , N_a , and N_b , the minimum delay required for achieving stability falls on the range of $[\text{round}(\bar{\eta}_{min}) - \Delta\eta, \text{round}(\bar{\eta}_{min}) + \Delta\eta]$. For a majority of the combinations where N_a and N_b are relatively small, we only need $\Delta\eta = 1$. As N_a , N_b and M become larger, the value of η_{min} could fall on longer range of consecutive values, and hence larger value of $\Delta\eta$ may be needed to increase the search span. All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ are list in Table 5.5, 5.6, and 5.7 for $M = 2, 4$, and 5 , respectively. Those combinations which are not listed only need $\Delta\eta = 1$.

Table 5.5 All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 2$. Timing mismatch case.

N_a	ε_{max}	N_b							
		3	4	5	6	7	8	9	10
19	0.09-0.1		2						
23	0.01								2
24	0.01								3
25	0.01	2						2	2
	0.02-0.1	2							
26	0.01					2	2	3	2
	0.02							2	
27	0.01	2	2		2	3	3	3	3
	0.02	2	2				2	2	2
	0.03-0.1	2	2						
28	0.01				2	3	2	3	5
	0.02							2	3
	0.03-0.04							2	2
29	0.01	2			2	3	4	5	5
	0.02	2					2	3	4
	0.03	2					2	3	3
	0.04	2						2	3
	0.05-0.07	2							2
	0.08-0.1	2							
30	0.01		2	2	2	3	5	6	5
	0.02		2		2	2	3	4	3
	0.03		2			2	2	3	3
	0.04		2				2	3	3
	0.05-0.07		2					2	2
	0.08, 0.1		2						
	0.09		2						2

Table 5.6 All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 4$. Timing mismatch case.

N_a	ε_{max}	N_b							
		3	4	5	6	7	8	9	10
19	0.01				2	2			
20	0.01					2	2		
21	0.01		2	2	2		2		
	0.02					2	2		
22	0.01		2	2	2	2	2	2	2
	0.02				2		2	2	
	0.03						2		
	0.09								2
23	0.01		2	3	3	3	2	3	2
	0.02		2	2	2	2	2	2	2
	0.03				2	2	2	2	2
	0.04-0.05				2	2	2	2	
24	0.01		2	3	2	3	3	2	3
	0.02		2	3	3	2	2	2	2
	0.03-0.05			2	3	2	2	2	2
	0.06-0.07				2	2	2	2	2
	0.08-0.09						2	2	2
25	0.01	2	3	4	3	4	3	2	3
	0.02		3	3	2	3	2	3	3
	0.03-0.05		2	3	2	2	2	2	2
	0.06				2	2		2	2
	0.07			2	2	2	2	2	2
	0.08-0.1					2		2	2
26	0.01	2	2	3	4	3	3	3	3
	0.02		2	3	3	4	3	3	3
	0.03		2	3	2	3	3	3	3
	0.04-0.05		2	2	2	3	3	3	2
	0.06-0.1			3	2	2	2	2	2
27	0.01	3	3	4	4	4	4	3	3
	0.02		2	3	4	3	4	3	4
	0.03-0.04		2	3	3	4	3	3	3
	0.05-0.07		2	3	3	3	3	2	3
	0.08-0.09			2	3	3	2	2	2
	0.1				2	3	2	3	2
28	0.01	4	3	4	5	4	3	4	4
	0.02	2	3	4	4	4	4	3	3
	0.03		3	4	4	3	4	3	4
	0.04-0.05		3	4	3	3	3	3	3
	0.06-0.07		3	3	3	3	3	3	3
	0.08-0.1		2	3	3	3	3	3	3
29	0.01	4	3	5	5	5	4	5	5
	0.02	2	2	4	5	4	4	4	4
	0.03		3	4	4	4	4	4	3
	0.04-0.05		3	4	4	3	4	3	3
	0.06-0.07		3	4	3	3	3	3	4
	0.08-0.09		3	4	3	3	3	3	3
	0.1		3	3	3	3	3	3	3
30	0.01	5	3	5	5	5	5	5	6
	0.02	4	3	4	4	5	4	5	4
	0.03	2	3	5	4	4	4	4	4
	0.04-0.05		3	4	4	4	4	4	3
	0.07-0.08		3	4	3	4	4	4	3
	0.06, 0.09, 0.1		2	4	3	4	4	3	3

Table 5.7 All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 5$. Timing mismatch case.

N_a	ε_{max}	N_b							
		3	4	5	6	7	8	9	10
21	0.01		2	2					
22	0.01		2		2		2		
	0.02		2	2					
	0.03			2					
23	0.01		2		2	2	2	2	2
	0.02-0.03		2		2				
	0.04-0.06					2			
24	0.01	2	2	2	3	2	2	3	2
	0.02		2		2		2		
	0.03		2		2				
	0.04-0.09				2				
25	0.01	2	3	3	3	3	3	3	3
	0.02		2	2	3	2	2	2	2
	0.03		2	2	2	2	2	2	
	0.04		2	2	2	2	2		
	0.05			2	2		2		
	0.06-0.1				2				
26	0.01	3	3	3	3	3	3	3	3
	0.02		3	2	2	3	2	3	3
	0.03		3	2	2	2	2	2	2
	0.04		3	2	3	2	2	2	2
	0.05-0.09		2	2	2	2	2	2	
	0.1			2	2	2	2		
27	0.01	3	3	2	4	3	4	4	3
	0.02	2	2	2	3	3	3	2	3
	0.03-0.04		2	2	2	3	2	3	3
	0.05-0.1		3	2	2	2	2	2	2
28	0.01	4	3	3	4	4	4	4	4
	0.02	2	3	2	4	4	4	3	4
	0.03-0.04		3	2	3	3	3	3	4
	0.05-0.6		2	2	3	3	2	3	3
	0.07-0.1		2		2	3	2	3	2
29	0.01	4	4	4	5	5	5	5	4
	0.02	3	3	3	4	4	4	4	4
	0.03	2	3	2	4	4	4	3	4
	0.04		3	2	3	4	3	3	4
	0.05-0.1		3	2	3	3	3	3	3
30	0.01	4	4	5	5	5	5	5	5
	0.02	4	4	3	5	4	5	5	4
	0.03	3	4	3	5	4	4	4	4
	0.04-0.05	2	3	3	4	4	4	3	4
	0.06	2	3	2	3	3	3	3	4
	0.07-0.08		3	2	3	4	3	3	4
	0.09		3	2	4	4	3	3	3
	0.1		3	2	3	4	3	2	3

5.4.2 Bandwidth Mismatch Case

Assume the bandwidth of the S/H circuit in Channel- m , $m = 1, 2, \dots, M - 1$ is within the range of $(1 - \varepsilon_{max}, 1 + \varepsilon_{max})\Omega_s$ where ε_{max} measures the scale of the bandwidth mismatch and Ω_s is the TIADC sampling frequency in radians per second. ADC-0 is chosen as the reference and its bandwidth is assumed to be Ω_s . In the experiment, the value of ε_{max} is varied from 0.05 to 0.3 with a step size of 0.05. For each value of ε_{max} , $(M - 1)$ random numbers are drawn from the uniform distribution over the interval of $(1 - \varepsilon_{max}, 1 + \varepsilon_{max})\Omega_s$. The $(M - 1)$ random numbers represent the bandwidths of ADC- m where $m = 1, 2, \dots, M - 1$. This process is repeated 10 thousand times so that a total of 10 thousand sets of bandwidth mismatches are generated for each value of ε_{max} . Then for each set of the bandwidth mismatch values, IIR filters are designed with various combinations of N_a and N_b . As in the timing mismatch case, the value of N_a is varied from 10 to 30 and the value of N_b is varied from 2 to 10 for each value of N_a . In each design, the value of η is varied from 0 to $N_a - 1$ to identify the minimum group delay required to achieve stability, i.e., η_{min} . The numbers of interleaving channels simulated are $M = 2, 4$, and 5. All the IIR filters are optimized in the least squares sense over the frequency range of $[0, 0.45f_s]$. Table 5.8 lists the design parameters and their values simulated in the experiment.

Table 5.8 List of the values of the design parameters simulated to identify η_{min} . Bandwidth mismatch case.

Optimization frequency range : $[0, 0.45f_s]$	
M	: 2, 4, 5
ε_{max}	: 0.05, 0.1, ..., 0.3
N_a	: 10, 11, ..., 20
N_b	: 2, 3, ..., 10

There are a total of 3402 different combinations of M , ε_{max} , N_a , and N_b simulated and for each combination, 10 thousand cases of η_{min} value are obtained. Similar to the timing mismatch case, we found that in more than 99% of the 10 thousand cases, the value of η_{min} fell on k consecutive numbers where k is ranging from one to nine for different combinations of M , ε_{max} , N_a , and N_b . The

occurrence of k in the 3402 combinations are counted and the histogram of k is plotted in Figure 5.11. From Figure 5.11, we can see that for a significant number of the design parameter combinations simulated, the value of η_{min} falls on a single number.

Figure 5.11 Occurrence of k in the 3402 different sets of design parameters where k is the range of consecutive numbers that more than 99% of the 10 thousand cases of η_{min} fall on for each set of design parameters. Bandwidth mismatch case.

We then wish to express $\bar{\eta}_{min}$ as a function of M , N_a , N_b and ϵ_{max} , as in (5.12) and then derive the range of numbers that η_{min} may fall on. As will be shown later, for most of the combinations of the design parameters, no more than three values of the group delay need to be tested to obtain η_{min} . The relationship between $\bar{\eta}_{min}$ and N_a is first investigated. Figure 5.12 shows the plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and ϵ_{max} while keeping the number of interleaving channels, M , fixed to 5. In each subplot, the value of N_b is fixed to a specific value and the 6 curves correspond to the cases of $\epsilon_{max} = 0.05, 0.1, \dots, 0.3$. $\bar{\eta}_{min}$ increases approximately linearly with N_a when only bandwidth mismatch is present. It is also observed that $\bar{\eta}_{min}$ increases with increasing ϵ_{max} . However, the effect of ϵ_{max} on $\bar{\eta}_{min}$ is noticeable only when $N_a \geq 25$ and $N_b \geq 5$ and hence the effect of ϵ_{max} on $\bar{\eta}_{min}$ is only secondary when $\epsilon_{max} \leq 0.3$. In Figure 5.13, the value of ϵ_{max} is fixed to 0.3, and $\bar{\eta}_{min}$ is plotted against N_a for various combinations of N_b and M . $\bar{\eta}_{min}$ and N_a have an approximately

linear relationship for $M = 2, 4$, and 5. $\bar{\eta}_{min}$ increases with M with N_a , N_b , and ε_{max} fixed.

Figure 5.12 Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and ε_{max} , M is fixed to 5. Bandwidth mismatch case.

Figure 5.13 Plots of $\bar{\eta}_{min}$ versus N_a for various combinations of N_b and M , ε_{max} is fixed to 0.3. Bandwidth mismatch case.

Figure 5.14 Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and ϵ_{max} , M is fixed to 4. Bandwidth mismatch case.

Figure 5.15 Plots of $\bar{\eta}_{min}$ versus N_b for various combinations of N_a and M , ϵ_{max} is fixed to 0.2. Bandwidth mismatch case.

Next we consider the effect of N_b on $\bar{\eta}_{min}$. The value of $\bar{\eta}_{min}$ is plotted against N_b in Figure 5.14 for different combinations of N_a and ε_{max} while keeping M fixed to 4. In each subplot, the value of N_a is fixed to a specific number and the value of ε_{max} is varied from 0.05 to 0.3. Clearly, $\bar{\eta}_{min}$ increases with N_b and the plotted curves are in the shapes of logarithmic functions. The effect of ε_{max} on $\bar{\eta}_{min}$ is negligible when $N_a < 26$. Figure 5.15 plots $\bar{\eta}_{min}$ versus N_b for different values of M while keeping ε_{max} fixed to 0.2. $\bar{\eta}_{min}$ increases with the number of interleaving channels.

Then, the effect of ε_{max} on $\bar{\eta}_{min}$ for various combinations of N_a , N_b and M is investigated. The values of $\bar{\eta}_{min}$ when ε_{max} is varied from 0.05 to 0.3 are plotted in Figure 5.16. The 21 lines in each plot correspond to $N_a = 10, 11, \dots, 30$. Figure 5.16(a) shows the case where $N_b = 3$ and $M = 2$. $\bar{\eta}_{min}$ remains constant regardless of the value of ε_{max} . Figure 5.16(b) shows the case where $N_b = 10$ and $M = 5$. We can see from Figure 5.16(b) that only when $N_a \geq 26$, the effect of ε_{max} to $\bar{\eta}_{min}$ becomes noticeable and $\bar{\eta}_{min}$ increases as ε_{max} increases. Hence the effect of ε_{max} on $\bar{\eta}_{min}$ is only secondary and we shall ignore it for now and express $\bar{\eta}_{min}$ as a function of N_a , N_b , and M only, i.e., $\bar{\eta}_{min} = f(M, N_a, N_b)$. The error caused by this simplification can be absorbed into the term $\Delta\eta$.

Figure 5.16 Plots of $\bar{\eta}_{min}$ versus ε_{max} for different values of N_a . Bandwidth mismatch case. (a) The case where $N_b = 3$ and $M = 2$. (b) The case where $N_b = 10$ and $M = 5$.

It is observed in Figure 5.12 and 5.13 that the value of $\bar{\eta}_{min}$ increases almost linearly with N_a .

Assume $\bar{\eta}_{min}$ and N_a have the following linear relationship

$$\bar{\eta}_{min}|_{N_b, M \text{ fixed}} \approx v_0 + v_1 N_a. \quad (5.17)$$

The coefficients v_0 and v_1 are then determined using line fitting method for each combination of N_b and M . Next, for each M , we attempt to express the coefficients $v_l, l = 0, 1$ as functions of N_b . It is observed in Figure 5.14 and 5.15 that $\bar{\eta}_{min}$ increases roughly logarithmically with N_b . Several different functions of N_b have been compared and we assume the coefficients $v_l, l = 0, 1$ are approximated using the following function of N_b ,

$$v_l|_{M \text{ fixed}} \approx \sum_{m=0}^2 v_{l,m} (\ln(\ln(N_b)))^m, \quad l = 0, 1. \quad (5.18)$$

The coefficients $v_{l,m}$ where $m = 0, 1, 2$ are determined using curve fitting method for each M . To summarize, when considering only the bandwidth mismatch, the value of $\bar{\eta}_{min}$ for a specific number of interleaving channels, can be estimated using the following equation

$$\bar{\eta}_{min}|_{M \text{ fixed}} \approx v_{0,0} + v_{1,0} N_a + (v_{0,1} + v_{1,1} N_a) \ln(\ln(N_b)) + (v_{0,2} + v_{1,2} N_a) (\ln(\ln(N_b)))^2. \quad (5.19)$$

The coefficient values $v_{l,m}$ where $l = 0, 1$ and $m = 0, 1, 2$ for $M = 2, 4$ and 5 are listed in Table 5.9. Using (5.19) and the coefficients listed in Table 5.9, the RMS error in approximating $\bar{\eta}_{min}$ are 0.597, 0.380 and 0.316 for $M = 2, 4$ and 5 respectively in the bandwidth mismatch case.

Finally, we need to identify the values of $\Delta\eta$ such that for each combination of M , ε_{max} , N_a , and N_b , the minimum delay required for achieving stability falls on the range of $[\text{round}(\bar{\eta}_{min}) - \Delta\eta, \text{round}(\bar{\eta}_{min}) + \Delta\eta]$. When N_a or N_b are relatively small, we only need $\Delta\eta = 1$, meaning that the range to search for η_{min} consists of only three numbers. As N_a , N_b and M become larger, the value of η_{min} could fall on longer span of consecutive numbers, and hence larger value of $\Delta\eta$ is needed. Table 5.10, 5.11 and 5.12 lists all the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 2, 4$ and 5 , respectively. Those combinations which are not listed only need $\Delta\eta = 1$.

Table 5.9 The coefficients in (5.19) for approximate $\bar{\eta}_{min}$. Bandwidth mismatch case.

	$v_{0,0}$	$v_{1,0}$	$v_{0,1}$	$v_{1,1}$	$v_{0,2}$	$v_{1,2}$
$M = 2$	-1.327	0.530	0.156	0.168	0.948	0.170
$M = 4$	-0.648	0.619	0.399	0.303	-0.755	-0.006
$M = 5$	-0.677	0.638	0.650	0.291	-1.290	0.004

Table 5.10 All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 2$. Bandwidth mismatch case.

N_a	ε_{max}	N_b							
		3	4	5	6	7	8	9	10
21	0.05, 0.2, 0.25, 0.3		2						
22	0.05								2
23	0.05							2	2
24	0.05		2				2	3	2
25	0.05					2	2	2	2
	0.1							2	2
26	0.05				2	3	3	3	4
	0.1						2	2	2
	0.15							2	2
27	0.05	2	2		3	3	3	4	5
	0.1					2	2	3	3
	0.15						2	2	3
	0.2							2	2
	0.25								2
28	0.05				2	3	4	5	5
	0.1					2	2	3	3
	0.15, 0.2						2	3	3
	0.25, 0.3							2	3
29	0.05	2	2	2	2	3	5	5	5
	0.1	2	2		2	3	3	4	4
	0.15	2	2		2	2	3	4	3
	0.2	2	2			2	2	3	3
	0.25, 0.3	2	2				2	3	3
30	0.05		2	3	3	4	5	6	6
	0.1				2	3	4	4	5
	0.15, 0.2				2	3	3	3	4
	0.25					2	3	3	3
	0.3						3	2	3

Table 5.11 All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 4$. Bandwidth mismatch case.

N_a	ε_{max}	N_b							
		3	4	5	6	7	8	9	10
18	0.05					2			
19	0.05			2	2	2	2		
20	0.05		2	2	2		2	2	
	0.1			2		2	2		
21	0.05		2	2	2	2		2	2
	0.1				2		2	2	
	0.15						2	2	
	0.2						2		
22	0.05		2	3	3	3	2	3	2
	0.1		2	2	2	2		2	2
	0.15, 0.2				2	2	2	2	2
	0.25						2	2	
	0.3						2		
23	0.05		2	3	2	3	3	2	3
	0.1		2	3	2	2	2	3	2
	0.15		2	2	3	2	2	2	2
	0.2			2	3	2		2	2
	0.25, 0.3				2	2		2	2
24	0.05	3	2	3	3	3	3	3	2
	0.1, 0.15		3	3	2	3	2	2	3
	0.2		2	3	2	2	2	2	2
	0.25			2	2	2	2	2	2
	0.3			2	2	2		2	2
25	0.05	2	2	3	4	3	4	3	3
	0.1, 0.15		2	3	3	3	3	2	3
	0.2, 0.25		2	2	2	3	3	2	3
	0.3			2	2	3	2	2	2
26	0.05	3	3	4	4	4	3	4	4
	0.1		2	3	4	3	4	3	3
	0.15		2	3	3	3	3	3	4
	0.2, 0.25, 0.3		2	3	3	3	3	2	3
27	0.05	5	3	4	4	4	4	5	5
	0.1	2	3	4	4	3	4	4	3
	0.15		3	4	4	3	4	3	3
	0.2		3	4	3	3	3	3	2
	0.25, 0.3		3	3	3	3	3	3	4
28	0.05	4	3	4	5	5	5	5	5
	0.1	3	2	5	4	4	4	4	4
	0.15, 0.2	2	2	4	4	4	4	4	3
	0.25, 0.3		2	4	4	3	4	4	3
29	0.05	5	3	5	5	5	5	6	5
	0.1	4	3	4	5	5	4	5	5
	0.15	2	3	4	4	4	4	4	4
	0.2	2	2	4	4	4	3	4	3
	0.25, 0.3		2	3	4	4	4	4	3
30	0.05	5	4	5	6	6	6	6	6
	0.1	5	3	5	5	4	5	5	6
	0.15	4	3	4	5	5	5	5	5
	0.2	3	3	4	5	5	4	5	4
	0.25	2	3	4	5	5	4	4	4
	0.3	2	3	4	4	4	4	4	4

Table 5.12 All the combinations of N_a , N_b and ε_{max} which need $\Delta\eta \geq 2$ for $M = 5$. Bandwidth mismatch case.

N_a	ε_{max}	N_b							
		3	4	5	6	7	8	9	10
18	0.05		2						
20	0.05		2	2					
21	0.05		2		2				
	0.1		2	2					
	0.15, 0.2			2					
22	0.05		2		2	2	2	2	2
	0.1		2		2		2		
	0.15				2				
	0.2		2		2				
	0.3					2			
23	0.05	2	2	2	2	2	2	3	2
	0.1		2		2	2	2	2	
	0.15, 0.2		2		2		2		
	0.25, 0.3				2				
24	0.05	2	3	2	3	3	3	2	3
	0.1, 0.15		2	2	2	2	2	2	2
	0.2		2	2	2	2			
	0.25		2	2	2		2		
	0.3				2		2		
25	0.05	3	2	2	3	3	3	3	3
	0.1	2	3	3	3	3	2	2	3
	0.15, 0.2		3	2	2	2	2	2	2
	0.25		3	2	2	2		2	
	0.3		2	2		2		2	
26	0.05	3	3	3	4	3	4	4	3
	0.1	3	2	2	3	4	3	3	3
	0.15	2	2	2	3	3	2	2	3
	0.2		2		2	3	2	3	3
	0.25, 0.3		2	2	2	2	2	3	2
27	0.05	4	3	4	5	4	4	4	4
	0.1	2	3	3	4	3	4	4	3
	0.15	2	3	2	3	4	3	3	4
	0.2		3	2	3	3	3	3	3
	0.25, 0.3		2	2	3	3	2	3	3
28	0.05	4	4	4	4	5	4	5	4
	0.1	4	3	4	5	4	4	4	4
	0.15	3	3	3	4	3	4	3	3
	0.2	2	3	3	3	3	4	3	4
	0.25, 0.3		3	2	3	3	3	3	4
29	0.05	4	4	4	5	6	5	4	5
	0.1, 0.15	4	4	4	4	4	4	5	4
	0.2, 0.25	3	3	3	4	4	4	4	4
	0.3		3	2	4	3	4	4	4
30	0.05	5	4	5	6	5	5	5	4
	0.1	4	4	4	5	5	5	6	5
	0.15	4	3	4	4	5	4	5	4
	0.2	3	3	4	4	4	5	5	4
	0.25	2	3	3	4	4	4	4	4
	0.3		3	3	4	4	4	4	4

5.5 Chapter Summary

In this chapter, two “computer-aided” design approaches, namely, *Method-1*: by increasing the group delay η , and *Method-2*: by constraining the magnitudes of the denominator coefficients, have been introduced for gradually achieving a stable IIR compensation filter design. Theoretical proof of *Method-1* by using *Rouché’s Theorem* has been provided. It is found through numerous simulations that for fixed N_a , N_b and mismatch errors, increasing the group delay η reduces the maximum pole radius of the IIR compensation filter. For *Method-2*, as the denominator weights approach infinity, the magnitudes of the denominator coefficients approach zero and the IIR filter becomes an FIR filter. Thus, stability can always be achieved by adding sufficiently large weights to the denominator coefficients. In the design of the IIR compensation filter, we do not know the minimum group delay or the minimum denominator weights to achieve stability and thus a few iterations may be needed to obtain a stable IIR filter design.

To provide a general guideline for selecting a suitable value of η to achieve stability, we have derived explicit formulas expressing the minimum group delay needed to achieve stability, η_{min} , as a function of the various design parameters. Two situations have been investigated; one is the case where only timing mismatch is present and the other is the case where only bandwidth mismatch is present. The relationship between η_{min} and N_a is approximately linear whereas the relationship between η_{min} and N_b is in the shape of a logarithmic function. The effect of the scale of the mismatch, ε_{max} , on η_{min} is noticeable only when N_a , N_b and M are relatively large and hence the effect is only secondary. With the obtained expressions, the search span for η_{min} is narrowed from N_a down to $2\Delta\eta + 1$ where $\Delta\eta$ is ranging from 1 to 6 for different combinations of M , N_a , N_b and ε_{max} .

CHAPTER 6 : SIMULATION EXAMPLES AND DISCUSSION

6.1 Introduction

In this chapter, we shall use some examples to demonstrate the performance of our IIR compensation filter. We will show that (1) IIR filters with guaranteed stability for TIADC mismatch compensation can be designed and (2) our IIR filters have lower implementation complexity than the conventional FIR TIADC compensation filters in [111] when they are designed to achieve the same level of compensation performance, or much better compensation performance when they are designed with the same total number of filter coefficients. Other important aspects of IIR TIADC compensation filters such as the performance in the presence of identification error, filter redesign to cope with mismatch drifting, pipeline implementation etc will also be discussed. Throughout this chapter, we shall denote the IIR compensation filter whose stability is achieved by using *Method-1* as IIR-1 and that whose stability is achieved by using *Method-2* as IIR-2 for ease of illustration.

6.2 Timing and Bandwidth Mismatch Example

In this section, we consider both the timing and bandwidth mismatches and compare the compensation performance of our IIR filter to that of the conventional FIR filter in [111] when they are designed with the same total number of coefficients.

Example 6.1

Consider an example with $M = 5$, i.e., a TIADC constructed using five sub-ADCs. The timing and bandwidth mismatch parameters are shown in Table 6.1. ADC-0 is chosen as the reference and so it has no timing mismatch and its bandwidth is assumed to be Ω_s where $\Omega_s = 2\pi f_s$ and f_s is the overall sampling rate of the TIADC. Mismatch compensation filters are designed for Channel-1, 2, 3, and 4 to compensate for the relative mismatches.

Table 6.1 Mismatch parameters for the 5-channel TIADC in Example 6.1

$G_0(j\Omega) = \frac{1}{1 + j\Omega/\Omega_s}$	$G_1(j\Omega) = \frac{e^{-j0.22\Omega T_s}}{1 + j\Omega/(1.15\Omega_s)}$	$G_2(j\Omega) = \frac{e^{-j0.18\Omega T_s}}{1 + j\Omega/(0.93\Omega_s)}$
$G_3(j\Omega) = \frac{e^{j0.29\Omega T_s}}{1 + j\Omega/(0.75\Omega_s)}$	$G_4(j\Omega) = \frac{e^{j0.35\Omega T_s}}{1 + j\Omega/(1.32\Omega_s)}$	

In the simulation, we assume a multi-tone input signal consisting the sum of seven sinusoids whose frequencies are $0.0625if_s$, $i = 1, 2, \dots, 7$. The spectra lines of the seven sinusoids are shown in Figure 6.1(a). Figure 6.1(b) shows the frequency spectrum of the TIADC output without compensation. Both FIR and IIR filters are designed to compensate for the mismatches. The filters are optimized in the least squares sense over the frequency range from 0 to $0.4375f_s$. The least squares weighting function is assigned such that $w_m(\Omega) = 1$ for $0 \leq \Omega \leq 0.875\pi f_s$ and $w_m(\Omega) = 0$ for $\Omega > 0.875\pi f_s$. Figure 6.1(c) shows the frequency spectrum of the compensated output where the compensation filters are FIR filters consisting of a total of $4 \times 30 = 120$ coefficients. The value of η is 15 and the compensation performance in terms of SFDR is 60.4 dB. The coefficient values of the FIR filters are shown in Table 6.2.

For our proposed IIR compensation filters, each filter is allocated with 27 numerator coefficients and 3 nonzero denominator coefficients. The total number of coefficients is then $4 \times (27 + 3) = 120$, i.e., the same number as the FIR case. For stability achieved using *Method-1*, the value of η is 20 and the maximum pole radius, p_{max} , is 0.932. Denote the IIR compensation filter whose stability is achieved by using *Method-1* as IIR-1. Figure 6.1(d) shows the compensated output spectrum when using IIR-1.

The SFDR after compensation is 92.5 dB which is 32.1 dB higher than that achieved by the FIR filter. The coefficient values of IIR-1 are shown in Table 6.3. Note that in Table 6.3, the coefficients $b_1[1]$, $b_2[2]$, $b_3[3]$ and $b_4[4]$ equal to zero. We have $N_b = 4$ but the nonzero denominator coefficients for each channel is only three. The reason is as follows. ADC-0 is chosen as the reference channel and thus the columns of the signal matrix $\mathbf{S}_m$ corresponding to the compensated output of ADC-0 may be identical to the columns corresponding to the ADC-0 raw data. To ensure that the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ in (3.30) is well conditioned, the corresponding compensated output samples from ADC-0 should not be used. This has been explained earlier in Section 3.4.

If the value of η is fixed to 15, i.e., the same value as in the FIR filter case, the IIR filter is not stable and the magnitudes of the denominator coefficients need to be constrained. Using *Method-2*, we impose a uniform weight to all the denominator coefficients. The uniform weight of 4.554×10^{-8} is selected by a trial and error “computer aided” design approach and a stable design with $p_{max} = 0.970$ is achieved. Several iterations may be needed to achieve a stable design. In each iteration, (5.8) through (5.11) are evaluated on a dense grid of frequencies from 0 to $0.4375f_s$ to obtain the coefficient vectors $\mathbf{h}_m$ where $m = 1, 2, 3, 4$ and the value of p_{max} is calculated. In the next iteration, the denominator weight is adjusted according to the value of p_{max} . This process is repeated until a desired value of p_{max} is reached. Denote the IIR compensation filter whose stability is achieved by using *Method-2* as IIR-2. The frequency spectrum of the compensated output using IIR-2 is shown in Figure 6.1(e) and the SFDR is 82.1 dB. The coefficient values of IIR-2 are listed in Table 6.4. The frequency response of the IIR filters and hence the compensation performance is affected by a larger amount when greater weights are imposed on the denominator coefficients. To illustrate the effect, a uniform weight of 1.198×10^{-7} is imposed on all the denominator coefficients such that the maximum pole radius of the IIR filter is 0.930. The compensated output signal spectrum is shown in Figure 6.1(f). It is observed that the highest spurious components after compensation is -77.7 dB and is higher than that in Figure 6.1(e) which is -82.1 dB. The coefficient values of IIR-2 with $p_{max} = 0.930$ are shown in Table 6.5.

Figure 6.1 Simulated frequency spectrums in Example 6.1. (a) Input. (b) Uncompensated output. (c) Compensated output using FIR filter. (d) Compensated output using IIR-1 filter with $p_{max} = 0.932$. (e) Compensated output using IIR-2 filter with $p_{max} = 0.970$. (f) Compensated output using IIR-2 filter with $p_{max} = 0.930$. The total number of coefficients for (c) through (f) is each $4 \times 30 = 120$.

As can be seen from Figure 6.1, the SFDR after compensation when using the IIR compensation filter instead of the FIR filter can be improved by as much as 32 dB if both filters are designed with a total number of 120 coefficients. In this example, the performance of the IIR filter whose stability is achieved using *Method-1* is better than that whose stability is achieved using *Method-2*. Thus, *Method-2* is adopted only if the value of η is pre-assigned and that the obtained filter is not stable for the pre-assigned value of η . Using the FIR filter, the SFDR of the compensated output is 60.4 dB. If the IIR filter is designed to achieve the same level of SFDR, only 4×19 coefficients would be needed. Each IIR filter is allocated 16 numerator coefficients and 3 nonzero denominator coefficients. The value of η is 11 which is even smaller than that in the FIR case. The total number of multipliers used in our IIR compensation filter is less than $2/3$ of that used in the conventional FIR compensation filter [111].

Table 6.2 Coefficient values for the FIR compensation filter in Example 6.1

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$h_m[0]$	0	0.00049451	0.00059815	-0.00039315	-0.00035087
$h_m[1]$	0	-0.00142207	-0.00214919	0.00131446	0.00057180
$h_m[2]$	0	0.00162816	0.00503126	-0.00447878	-0.00168845
$h_m[3]$	0	-0.00129566	-0.00523199	0.01011313	0.00550863
$h_m[4]$	0	0.00269114	0.00369446	-0.01026302	-0.01204791
$h_m[5]$	0	-0.00757053	-0.00703349	0.00700364	0.01197721
$h_m[6]$	0	0.01517805	0.01893806	-0.01293123	-0.00790918
$h_m[7]$	0	-0.01428058	-0.03702365	0.03435089	0.01419873
$h_m[8]$	0	0.00871562	0.03430234	-0.06671939	-0.03717707
$h_m[9]$	0	-0.01474017	-0.02054574	0.06170124	0.07175094
$h_m[10]$	0	0.03795693	0.03461311	-0.03713796	-0.06625019
$h_m[11]$	0	-0.07378470	-0.09064609	0.06454123	0.04021165
$h_m[12]$	0	0.06973766	0.18207238	-0.17881423	-0.07329846
$h_m[13]$	0	-0.04734999	-0.18044593	0.39269700	0.22487390
$h_m[14]$	0	0.14711971	0.15110398	-0.44606030	-0.60952507
$h_m[15]$	1	1.06798347	1.08105072	1.07111307	1.21232243
$h_m[16]$	0	-0.31439336	-0.31528759	0.21146735	0.27448699
$h_m[17]$	0	0.14716055	0.27055484	-0.21277635	-0.09768022
$h_m[18]$	0	-0.03318022	-0.15126181	0.22295952	0.11350268
$h_m[19]$	0	0.02125562	0.03945798	-0.13382049	-0.12479863
$h_m[20]$	0	-0.02796613	-0.02674774	0.03730483	0.07615109
$h_m[21]$	0	0.03198324	0.03567634	-0.02583153	-0.02133694
$h_m[22]$	0	-0.01972658	-0.04086570	0.03432564	0.01443533
$h_m[23]$	0	0.00547753	0.02512938	-0.03885331	-0.01855344
$h_m[24]$	0	-0.00352822	-0.00688879	0.02356402	0.02026860
$h_m[25]$	0	0.00427183	0.00431033	-0.00621024	-0.01187755
$h_m[26]$	0	-0.00438588	-0.00506920	0.00361049	0.00287251
$h_m[27]$	0	0.00242382	0.00504825	-0.00388313	-0.00139930
$h_m[28]$	0	-0.00050746	-0.00271787	0.00347358	0.00118244
$h_m[29]$	0	0.00017212	0.00053795	-0.00164406	-0.00067879

Table 6.3 Coefficient values for the IIR-1 filter with $p_{max} = 0.932$ in Example 6.1

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$a_m[0]$	0	-0.00001032	-0.00000593	0.00000255	0.00000111
$a_m[1]$	0	0.00003355	0.00002498	-0.00001057	-0.00000224
$a_m[2]$	0	-0.00004134	-0.00006418	0.00004094	0.00000802
$a_m[3]$	0	0.00003683	0.00007113	-0.00010059	-0.00002968
$a_m[4]$	0	-0.00008605	-0.00005582	0.00010845	0.00007092
$a_m[5]$	0	0.00026151	0.00011922	-0.00008210	-0.00007526
$a_m[6]$	0	-0.00055442	-0.00034676	0.00017033	0.00005576
$a_m[7]$	0	0.00054342	0.00071714	-0.00048951	-0.00011425
$a_m[8]$	0	-0.00035561	-0.00069237	0.00100681	0.00032752
$a_m[9]$	0	0.00064488	0.00044441	-0.00097031	-0.00067514
$a_m[10]$	0	-0.00171823	-0.00079799	0.00062408	0.00065341
$a_m[11]$	0	0.00337203	0.00213189	-0.00113610	-0.00042621
$a_m[12]$	0	-0.00315604	-0.00421925	0.00309400	0.00080009
$a_m[13]$	0	0.00195615	0.00399334	-0.00625983	-0.00225607
$a_m[14]$	0	-0.00353460	-0.00255957	0.00606051	0.00473668
$a_m[15]$	0	0.00992121	0.00498432	-0.00412065	-0.00475873
$a_m[16]$	0	-0.02119779	-0.01535096	0.00906611	0.00355729
$a_m[17]$	0	0.02202323	0.03665900	-0.03257434	-0.00957926
$a_m[18]$	0	-0.01895748	-0.04328380	0.09484869	0.04551801
$a_m[19]$	0	0.09477189	0.05741959	-0.14502205	-0.20355765
$a_m[20]$	1	1.23794552	1.38469405	0.79930449	0.72987876
$a_m[21]$	0	-0.82823120	2.10931968	-2.63336514	1.69128605
$a_m[22]$	0	1.37773990	-0.61864366	-1.49532128	-1.06381828
$a_m[23]$	0	1.04545613	0.59609226	0.19956451	-0.25642999
$a_m[24]$	0	-0.08456556	0.16669308	-0.09803533	0.01338007
$a_m[25]$	0	0.00260507	0.00058581	-0.00251774	-0.00223578
$a_m[26]$	0	-0.00024730	-0.00007807	0.00005990	0.00005171
$b_m[1]$	0	0	2.20665269	-2.37717608	1.46164877
$b_m[2]$	0	1.25362208	0	-1.89343780	-1.12389507
$b_m[3]$	0	0.70389758	0.38305933	0	-0.39138783
$b_m[4]$	0	-0.12286107	0.08814185	-0.03544248	0

Table 6.4 Coefficient values for the IIR-2 filter with $p_{max} = 0.970$ in Example 6.1

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$a_m[0]$	0	-0.00013008	-0.00054376	0.00063274	0.00032723
$a_m[1]$	0	0.00031170	0.00042138	-0.00068472	-0.00077174
$a_m[2]$	0	-0.00097478	-0.00090120	0.00052670	0.00081414
$a_m[3]$	0	0.00212964	0.00265783	-0.00113722	-0.00060289
$a_m[4]$	0	-0.00214832	-0.00560629	0.00342149	0.00126229
$a_m[5]$	0	0.00150564	0.00553369	-0.00738874	-0.00375290
$a_m[6]$	0	-0.00308400	-0.00377278	0.00747196	0.00809263
$a_m[7]$	0	0.00933815	0.00767994	-0.00541817	-0.00822457
$a_m[8]$	0	-0.02096403	-0.02382095	0.01247120	0.00613913
$a_m[9]$	0	0.02248470	0.05608351	-0.04489657	-0.01565095
$a_m[10]$	0	-0.02013844	-0.06441217	0.12743482	0.06773755
$a_m[11]$	0	0.10007985	0.07682036	-0.18617232	-0.26789966
$a_m[12]$	0	1.20765960	1.28375715	0.85937655	0.82641044
$a_m[13]$	0	-0.68245760	0.80068856	-1.09357811	1.19705357
$a_m[14]$	0	0.84560149	0.22208921	0.80091055	0.22617324
$a_m[15]$	0	-0.10167433	-0.71236142	-0.41039065	0.26793317
$a_m[16]$	0	0.30087822	-0.59899242	0.39031372	-0.02538221
$a_m[17]$	0	-0.03159490	-0.01320249	0.03068675	0.00497155
$a_m[18]$	0	0.01308223	0.00488741	-0.00361302	0.00010417
$a_m[19]$	0	-0.00436143	-0.00208961	0.00181856	-0.00025010
$a_m[20]$	1	0.00040611	0.00057320	-0.00099705	0.00031791
$a_m[21]$	0	-0.00007156	-0.00000477	0.00034089	-0.00030823
$a_m[22]$	0	0.00002387	-0.00001720	-0.00002247	0.00015796
$a_m[23]$	0	-0.00000374	0.00001330	-0.00000086	-0.00002666
$a_m[24]$	0	-0.00013008	-0.00054376	0.00063274	0.00032723
$a_m[25]$	0	0.00031170	0.00042138	-0.00068472	-0.00077174
$a_m[26]$	0	-0.00097478	-0.00090120	0.00052670	0.00081414
$b_m[1]$	0	0	1.14443418	-1.26874982	0.85964773
$b_m[2]$	0	0.53052509	0	0.53531557	0.08986381
$b_m[3]$	0	-0.24631746	-0.73274413	0	0.33504365
$b_m[4]$	0	0.35174941	-0.37587256	0.21433847	0

Table 6.5 Coefficient values for the IIR-2 filter with $p_{max} = 0.930$ in Example 6.1

	$m = 0$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$a_m[0]$	0	0.00005584	0.00008310	-0.00002878	-0.00002324
$a_m[1]$	0	-0.00017512	-0.00032436	0.00010930	0.00004262
$a_m[2]$	0	0.00021219	0.00080116	-0.00040699	-0.00014185
$a_m[3]$	0	-0.00018553	-0.00086639	0.00097905	0.00050374
$a_m[4]$	0	0.00043127	0.00065684	-0.00104348	-0.00117517
$a_m[5]$	0	-0.00131972	-0.00136679	0.00078063	0.00122896
$a_m[6]$	0	0.00283394	0.00394585	-0.00162704	-0.00089453
$a_m[7]$	0	-0.00281990	-0.00817716	0.00476452	0.00182879
$a_m[8]$	0	0.00192486	0.00795383	-0.01006218	-0.00532887
$a_m[9]$	0	-0.00380793	-0.00526446	0.00999126	0.01128799
$a_m[10]$	0	0.01121441	0.01027071	-0.00698590	-0.01129453
$a_m[11]$	0	-0.02459323	-0.03069966	0.01528189	0.00814996
$a_m[12]$	0	0.02586417	0.06987575	-0.05270317	-0.01965836
$a_m[13]$	0	-0.02225813	-0.07783358	0.14402351	0.08045304
$a_m[14]$	0	0.10426781	0.08656195	-0.20348758	-0.29997183
$a_m[15]$	0	1.19347244	1.25012812	0.87568531	0.86831086
$a_m[16]$	0	-0.63943284	0.54080235	-0.95632840	1.05038448
$a_m[17]$	0	0.74935465	0.28646549	0.69939090	0.24433597
$a_m[18]$	0	-0.11848430	-0.67295822	-0.31026164	0.09592453
$a_m[19]$	0	0.22271401	-0.40988271	0.26242575	0.01047628
$a_m[20]$	1	-0.01978476	-0.00200712	0.01519689	-0.00812314
$a_m[21]$	0	0.00668196	-0.00180761	-0.00073769	0.00194675
$a_m[22]$	0	-0.00177833	0.00255764	-0.00025092	-0.00094344
$a_m[23]$	0	0.00004954	-0.00145922	0.00054815	0.00091985
$a_m[24]$	0	0.00003417	0.00028827	-0.00034928	-0.00079105
$a_m[25]$	0	-0.00003917	-0.00009507	0.00007516	0.00037908
$a_m[26]$	0	0.00002270	0.00004749	-0.00002233	-0.00005762
$b_m[1]$	0	0	0.90265965	-1.14338214	0.70389648
$b_m[2]$	0	0.44457407	0	0.49699799	0.16179100
$b_m[3]$	0	-0.22765461	-0.61718038	0	0.16209742
$b_m[4]$	0	0.26752478	-0.23780703	0.13135511	0

The proposed IIR compensation filter is an M -PTV filter whose z -transform at time instant n , $H_n(z)$, is given by (3.11) and $H_n(z) = H_{n+M}(z)$. Hence, there are a total of M LTI IIR filters in the proposed compensation system, each with frequency response of $H_m(e^{j\omega})$ where $m = 0, 1, \dots, M-1$, operating in a round-robin manner. The magnitude responses of the 5 LTI filters for the FIR and IIR compensation systems designed in this example are plotted in Figure 6.2. For the IIR filters, $H_m(e^{j\omega})$ is obtained by evaluating (3.11) at $z = e^{j\omega}$. For the FIR filter, $H_m(e^{j\omega})$ is obtained by using $H_m(e^{j\omega}) = \sum_{n=0}^{N-1} h_m[n] e^{-jn\omega}$ where the filter coefficients $h_m[n]$ are listed in Table 6.2. It is observed that in the FIR case, the magnitude responses of $H_m(e^{j\omega})$, $m = 1, 2, 3, 4$ deviates noticeably from 0 dB and this may be attributed to the large mismatch errors being simulated in this example. Normally, if the mismatches are small, the magnitude responses approximate unity for the FIR compensation filters. For the IIR compensation filters, the deviation of the magnitude response of $H_m(e^{j\omega})$ from 0 dB is even more significant.

It has been demonstrated in Chapter 4 that the proposed M -PTV IIR filter can be transformed into an equivalent filter bank structure consisting of M LTI IIR filters, $R_m(e^{j\omega})$, $m = 0, 1, \dots, M-1$. The z -transform of the equivalent LTI IIR filter is given in (4.47) and $R_m(e^{j\omega})$ is obtained by evaluating (4.47) at $z = e^{j\omega}$. Figure 6.3 plots the magnitude responses of the M LTI IIR filters, $R_m(e^{j\omega})$, $m = 0, 1, \dots, M-1$, in the equivalent filter bank structure. The magnitude responses are approximately piecewise constant over frequency intervals of $[-0.5f_s, -0.3f_s]$, $[-0.3f_s, -0.1f_s]$, $[-0.1f_s, 0.1f_s]$, $[0.1f_s, 0.3f_s]$ and $[0.3f_s, 0.5f_s]$. The length of each of the frequency intervals is $f_s/5$. The IIR filters have sharper transitions than the FIR filter.

Figure 6.2 Magnitude responses of $H_m(e^{j\omega})$, $m=0, 1, \dots, M-1$ for the filters designed in Example 6.1. (a) FIR filter. (b) IIR-1 filter with $p_{\max} = 0.932$. (c) IIR-2 filter with $p_{\max} = 0.970$. (d) IIR-2 filter with $p_{\max} = 0.930$.

Figure 6.3 Magnitude responses of the LTI IIR filters, $R_m(e^{j\omega})$, $m=0, 1, \dots, M-1$ in the equivalent filter bank representation for the filters designed in Example 6.1. (a) FIR filter. (b) IIR-1 filter with $p_{\max} = 0.932$. (c) IIR-2 filter with $p_{\max} = 0.970$. (d) IIR-2 filter with $p_{\max} = 0.930$.

6.3 Bandwidth Mismatch Example

In this section, we consider the bandwidth mismatch only and compare the implementation complexity of our IIR filter to that of the FIR filter when they are designed to achieve the same level of SFDR.

Example 6.2

To facilitate comparison, we shall use the same specifications in *Example-II* in [111] for designing FIR mismatch compensation filters. The TIADC consists of five channels and the bandwidth mismatch parameters are shown in Table 6.6. ADC-0 is chosen as the reference and its bandwidth is assumed to be Ω_s where $\Omega_s = 2\pi f_s$ and f_s is the overall sampling rate of the TIADC. Mismatch compensation filters are designed for Channel-1, 2, 3, and 4 to compensate for the relative mismatch error.

Table 6.6 Bandwidth mismatch parameters for the 5-channel TIADC in Example 6.2

$G_0(j\Omega) = \frac{1}{1 + j\Omega/\Omega_s}$	$G_1(j\Omega) = \frac{1}{1 + j1.07\Omega/\Omega_s}$	$G_2(j\Omega) = \frac{1}{1 + j1.15\Omega/\Omega_s}$
$G_3(j\Omega) = \frac{1}{1 + j0.93\Omega/\Omega_s}$	$G_4(j\Omega) = \frac{1}{1 + j0.95\Omega/\Omega_s}$	

We first consider a multi-tone input signal consisting of seven sinusoids whose frequencies are $0.0625if_s$, $i = 1, 2, \dots, 7$. The spectra lines of the seven sinusoids are shown in Figure 6.4(a) and the frequency spectrum of the TIADC output without compensation is shown in Figure 6.4(b). The complexity of our IIR filters is compared to that of the FIR filters when they are designed to achieve the same level of SFDR. To provide a more meaningful result, the mean squared errors (MSE) when using the FIR and the IIR filters are also compared. The MSE is defined as the mean squared value of the difference between the ideal and the estimated output samples, i.e.,

$$\text{MSE} = \frac{1}{P} \sum_{n=1}^P (\hat{x}[n] - \check{x}[n])^2, \quad (6.1)$$

where $\hat{x}[n]$ and $\check{x}[n]$ denote the estimated and the ideal samples, respectively, and P is the total number of samples used in the calculation. In this example, we used $P = 10000$ samples. Both the FIR and IIR filters are optimized in the least squares sense on 4376 discrete frequency points spaced evenly from 0 to $0.4375f_s$. $w_m(\Omega)$ is assigned the value one for the 4376 frequency points. For the FIR compensation filter in [111], forty-one coefficients for each channel are used. The value of η is 20. The frequency spectrum of the compensated output using the FIR filters is reproduced in Figure 6.4(c). The SFDR after compensation is 102.5 dB and the MSE is -101 dB.

We aim to design IIR compensation filters which achieve at least 102.5 dB SFDR with the least total number of coefficients while keeping $p_{max} \leq 0.95$. Both methods discussed in Chapter 5 for achieving stability are attempted. From our experience, for fixed N , where $N = N_a + N_b$ is the total number of filter coefficients for each channel, the compensation performance is the best when N_b is between 2 to M . For *Method-1*, the process of obtaining the IIR filter which satisfies all the above requirements is as follows.

1. From our experience, when the FIR and IIR filters are designed to achieve the same level of SFDR, the value of N required by the IIR filter is about 60% to 70% of that required by the FIR filter when using *Method-1*. Here we start with $N = 28$.
2. Vary the value of N_b from 2 to M . For each combination of N_a and N_b , use (5.19), Table 5.9 and Table 5.12 to predict the range of numbers that the minimum value of η required to achieve stability may fall on. For example, for $N_a = 24$ and $N_b = 4$, the value of η that achieves stability falls on the range between 15 to 19.
3. For each combination of N_a , N_b , and η , design the IIR filter using (3.25)-(3.28) and (3.30) and calculate the value of p_{max} . Stop if a design with at least 102.5 dB SFDR and $p_{max} \leq 0.95$ is found and store the combination of N_a , N_b , and η .
4. If a satisfactory design is found in Step 3, then decrease the value of N to 27 and start from

Step 1 again.

5. If no satisfactory design is found in Step 3, then increase the value of N to 29 and start from Step 1 again.
6. For all the designs obtained with at least 102.5 dB SFDR and $p_{max} \leq 0.95$, choose the one with the least total number of filter coefficients. Finally we have $N_a = 24$, $N_b = 4$, and $\eta = 18$ and the value of p_{max} is 0.896.

Denote the IIR filter whose stability is achieved by using *Method-1* as IIR-1. Figure 6.4(d) shows the compensated output spectrum using IIR-1. The SFDR after compensation is 103 dB and the MSE is -101.4 dB. For stability achieved using *Method-2*, we assume that the value of η is fixed at $\text{round}(N_a/2)$. Several combinations of N_a and N_b are tested and the IIR filter which satisfies all the requirements need 30 feed-forward coefficients and 2 feedback coefficients for each channel. A uniform weight of 1.6×10^{-9} for all the denominator coefficients is selected by a trial and error design approach to reduce the maximum pole radius to 0.95. Denote this filter as IIR-2. Figure 6.4(e) shows the compensated output spectrum using IIR-2. The SFDR after compensation is 104.4 dB and the MSE is -103.1 dB. In this example, IIR-1 requires fewer number of filter coefficients than IIR-2 to achieve about the same level of SFDR. In the design of IIR compensation filters, *Method-2* is adopted only when the value of η is pre-assigned and that the IIR filter with the pre-assigned value of η do not fulfill the requirement of p_{max} .

Next, the compensation performances of the FIR and IIR filters when the input is a bandlimited Gaussian noise are investigated. A white Gaussian noise (WGN) with zero mean and variance $\sigma^2 = 1$ is filtered by a low pass filter with passband spanning from 0 to $0.4375f_s$ to generate the bandlimited input signal. The frequency spectrum of the TIADC output with bandlimited Gaussian noise input is shown in Figure 6.5(a). Since the input signal bandwidth is bandlimited to within $0.4375f_s$, the TIADC output signal will also be bandlimited if it has no mismatch. It can be seen from Figure 6.5(a) that the uncompensated TIADC output has considerably high signal power at frequencies higher than $0.4375f_s$; the out-of-band spurious signal is caused by the TIADC mismatches. The spectrum of the compensated output when using the FIR filter, IIR-1 and IIR-2 are shown in Figure 6.5(b), (c) and (d),

respectively. The out-of-band spurious signal has been suppressed by over 40 dB by the mismatch compensation filters. The MSE of the FIR filter, IIR-1 and IIR-2 are -103 dB, -103 dB and -103.3 dB, respectively, with bandlimited Gaussian noise input.

Figure 6.4 Simulated frequency spectrum in Example 6.2, assuming multi-tone input. (a) Input. (b) Uncompensated output. (c) Compensated output using FIR filter. (d) Compensated output using IIR-1 filter with $p_{max} = 0.896$. (e) Compensated output using IIR-2 filter with $p_{max} = 0.950$.

Figure 6.5 Simulated frequency spectrums in Example 6.2, assuming bandlimited Gaussian noise input. (a) Uncompensated output. (b) Compensated output using the FIR filter. (c) Compensated output using IIR-1 filter with $p_{max} = 0.896$. (d) Compensated output using IIR-2 filter with $p_{max} = 0.950$.

The implementation complexity and the performance of the FIR and IIR filters when the input is a multi-tone signal are summarized in Table 6.7. Note that for IIR-1, we have $N_a + N_b = 28$, but the computation load for estimating the output samples is only 27 multiplications per output sample. For IIR-2, we have $N_a + N_b = 32$, but the computation load for estimating the output samples are 31, 31, 32, and 32 multiplications per output sample for Channels 1, 2, 3, and 4, respectively. The reason has been explained in Section 3.4 and Example 6.1. From Table 6.7, if IIR-1 and IIR-2 are used instead of the FIR filter, 34.1% and 23.2% of the multipliers can be saved, respectively. The total numbers of the delay elements required by the two IIR filters are also much smaller than that in the FIR filter case.

Table 6.7 Comparison of the implementation complexity and performance of the FIR and IIR filters in Example 6.2.

	FIR[98]	IIR-1	IIR-2
Delay elements	40	27	31
Multipliers	164	108	126
Multiplications per sample	41	27	31 or 32
Overall delay in T_s	20	18	15
SFDR (dB)	102.5	103.0	104.4
MSE(dB)	-101.0	-101.4	-103.1

We have discussed in Section 3.4 that residue error is introduced when we use the corrected output samples rather than the ideal samples in the estimation process. If the FIR filter in [111] and our IIR filter are designed to achieve the same compensation performance, then the WLS approximation error of the IIR filter need to be lower than that of the FIR filter. Lower WLS error is required to counteract the effect of the residue error such that the actual compensation performance is the same as that of the FIR filter. Take the FIR and IIR-1 filters for example. The SFDR after compensation for both filters are 103 dB. In (5.6), we have defined the RMS value of the WLS error as $e_{rms} = \sqrt{\sum_{m=0}^{M-1} \sum_{l=1}^L |\varepsilon_m(j\Omega_l)|^2 / (ML)}$. The values of e_{rms}^2 for the FIR and the IIR-1 filter are -117.7 dB and -125.7 dB, respectively. Now suppose the FIR and the IIR-1 filters are designed to achieve 60

dB SFDR (larger residue error). The design parameters for the FIR filter are $N_a = 16$, $N_b = 0$, and $\eta = 7$ and that for the IIR-1 filter are $N_a = 9$, $N_b = 2$, and $\eta = 4$. The values of e_{rms}^2 for the FIR and IIR-1 filters are -70.9 and -74.9 dB, respectively. For the 60 dB SFDR cases, the saving of the number of multipliers is 34.4% which is similar to that in the 100 dB SFDR case. Hence the magnitude of the residue error has no obvious effect on the performance of IIR filter in this example.

As mentioned in Section 1.1, the availability of high-speed TIADC enables direct RF sampling of radar signals, eliminating the need for analog down-conversion stages. Currently, the fastest single-core ADC available in the market is EV10AS152A from e2v. It has 3 GS/s sampling rate, 10-bit resolution and 5 GHz analog input bandwidth. With this ADC, direct RF sampling of radar signal with frequency up to 5 GHz with 3GS/s sampling rate is possible. The EV10AS152A ADC is able to sample the radar signal operating in the C band (4 GHz to 8 GHz frequency band). If five such ADCs are time interleaved, the effective sampling rate can be increased to 15 GS/s but the analog input bandwidth remains 5 GHz. Although the Nyquist band is extended to 7.5 GHz, signals with frequency higher than 5 GHz are still attenuated by the front-end circuitry of the sub-ADCs. To overcome the bandwidth limitation, photonic sampling can be employed to extend the effective bandwidth to five times of the sub-ADC bandwidth, i.e., 25 GHz [136]. The 5-channel TIADC is able to sample up to the radar K band (18 GHz to 26.5 GHz frequency band) with 15 GS/s sampling rate. Assume the mismatch characteristics of the five sub-ADCs are those specified in Table 6.6. To achieve 102 dB SFDR, the FIR compensation filter require 41 multiplications to produce one output sample whereas the IIR-1 filter only require 27 multiplications per output sample. The DSP performance of an FPGA is usually measured by the number of multiply-accumulate operations per second. The state-of-the-art FPGA is able to perform thousands of Giga-multiply-accumulate operations per second (GMAC/s). For example, the Virtex Ultrascale+ family has 2280 to 11904 DSP slices and is able to perform 21213 GMAC/s. Implementation of the FIR and IIR compensation filters on the FPGA is possible. Compared to the FIR filter, implementing the IIR filter requires 34% less of DSP slices and hence the FPGA resources are freed to realize other functionalities.

6.4 Timing Mismatch Example

In this section, we consider the case where only timing mismatch is present. The coefficient saving when the IIR filters are used instead of the FIR filters in [111] to achieve the same level of SFDR is investigated for a range of timing mismatch errors. The numerical robustness of the FIR and the IIR filters are also investigated.

Example 6.3

In this example we consider the timing mismatch in a 5-channel TIADC. The coefficient saving when using our IIR compensation filter instead of the conventional FIR compensation filter is investigated for a range of mismatched timing. The analog transfer functions which models the timing mismatch in Channel- m is $G_m(j\Omega) = e^{j\Delta r_m \beta \Omega T_s}$ where $\beta \Delta r_m$ represents the timing offset in Channel- m normalized with respect to T_s and β is a positive constant that controls the scale of the timing mismatch. The values of Δr_m for the five channels are shown in Table 6.8. ADC-0 is chosen as the reference and thus $\Delta r_0 = 0$ and $G_0(j\Omega) = 1$.

Table 6.8 Timing mismatch parameters for the 5-channel TIADC in Example 6.3

$G_0(j\Omega) = 1$	$G_1(j\Omega) = e^{j0.36\beta\Omega T_s}$	$G_2(j\Omega) = e^{-j0.23\beta\Omega T_s}$
$G_3(j\Omega) = e^{j0.27\beta\Omega T_s}$	$G_4(j\Omega) = e^{-j0.32\beta\Omega T_s}$	

Assume the same multi-tone input signal which consists of seven sinusoids with frequencies $0.0625if_s$, $i = 1, 2, \dots, 7$. Figure 6.6(a) shows the frequency spectrum of the TIADC output without compensation for $\beta = 1$. Both the FIR and IIR filters are optimized in the least squares sense over the frequency range of $[0, 0.4375f_s]$. Suppose 100 dB SFDR is to be achieved after mismatch compensation. Forty-seven coefficients are required for each channel if FIR filters are used. The value of η is 23. The frequency spectrum of the compensated output using the FIR filter is shown in Figure 6.6(b) for $\beta = 1$. For the design of the IIR compensation filter, *Method-1* is used to constrain p_{max} to

be less than 0.95. The design parameters are $N_a = 30$, $N_b = 4$ and $\eta = 22$. The value of p_{max} is 0.949 and the SFDR after mismatch correction is 103.2 dB. Figure 6.6(c) shows the frequency spectrum of the compensated output using the IIR filter for $\beta = 1$. Note that the denominator coefficients $b_1[1]$, $b_2[2]$, $b_3[3]$, and $b_4[4]$ are zero because ADC-0 is chosen as the reference. The IIR filter only needs 33 coefficients for each channel and a saving of 29.8% of the total number of multipliers is obtained when using the IIR filter instead of the FIR filter for $\beta = 1$.

Figure 6.6 Simulated frequency spectra for $\beta = 1$ in Example 6.3. (a) Uncompensated output. (b) Compensated output using FIR filter; 4×47 coefficients are required for 100 dB SFDR. (c) Compensated output using IIR filter; only 4×33 coefficients are required and the SFDR is 103.2 dB.

We have also investigated the coefficient saving as the mismatch errors become smaller. To do this, the value of β is decreased gradually from unity to 10^{-3} in log scale with a step size of $10^{-0.25}$. For each value of β , FIR filters are designed to achieve 100 dB SFDR and then IIR filters are designed such that the SFDR after compensation is at least that achieved by the FIR filters. We aim to find the IIR filter design which uses the minimum total number of filter coefficients and has p_{max} less than 0.95. *Method-1* is used to constrain the maximum pole radius. The design parameters of the FIR and

IIR filters and the percentage savings of the total number of coefficients achieved by the IIR filters for different values of β are listed in Table 6.9. The number of coefficients for both the FIR and IIR compensation filters decrease as the scale of the timing mismatch decreases. The number of coefficients required by the IIR filters are always about 30% less than that required by the FIR filters.

Table 6.9 Coefficient savings in percentage for various values of β

β		N_a	N_b	η	p_{max}	Saving
1	FIR	47	0	23		29.8%
	IIR	30	4	22	0.949	
$10^{-0.25}$	FIR	46	0	22		30.4%
	IIR	28	5	22	0.928	
$10^{-0.5}$	FIR	44	0	21		31.8%
	IIR	25	6	20	0.937	
$10^{-0.75}$	FIR	42	0	21		31.0%
	IIR	26	4	19	0.949	
0.1	FIR	39	0	19		33.3%
	IIR	23	4	17	0.931	
$10^{-1.25}$	FIR	37	0	18		32.4%
	IIR	22	4	16	0.945	
$10^{-1.5}$	FIR	35	0	17		31.4%
	IIR	21	4	16	0.886	
$10^{-1.75}$	FIR	33	0	16		33.3%
	IIR	19	4	14	0.916	
0.01	FIR	30	0	14		35.8%
	IIR	15	5	11	0.947	
$10^{-2.25}$	FIR	28	0	14		31.3%
	IIR	15	5	11	0.947	
$10^{-2.5}$	FIR	25	0	12		35.0%
	IIR	14	3	9	0.932	
$10^{-2.75}$	FIR	23	0	11		29.3%
	IIR	14	3	9	0.932	
0.001	FIR	20	0	10		33.8%
	IIR	11	3	7	0.898	

For $\beta = 1$, the condition numbers of the matrix to be inverted, $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$, are 2.24×10^7 , 1.96×10^7 , 1.95×10^7 , and 2.25×10^7 for $m = 1, 2, 3$, and 4 , respectively, in the FIR case. For the IIR filter, the condition numbers are 3.20×10^{11} , 7.74×10^{10} , 4.97×10^{11} , and 1.07×10^{12} , respectively. Thus, the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filter is more ill-conditioned than the corresponding matrix in

the FIR case. The condition numbers of $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the FIR and IIR filters in Table 6.9 are listed in Table 6.10 and 6.11, respectively. In Table 6.9, the filters are designed to achieve 100 dB SFDR and thus as β decreases, the number of coefficients decrease. We can see from Table 6.10 and 6.11 that as β decreases, the condition numbers for the FIR filters decrease whereas the conditional numbers for the IIR filters remain about the same order of magnitude.

Table 6.10 Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the FIR filters in Table 6.9 where the filters are designed with fixed SFDR

β	$m = 1$	$m = 2$	$m = 3$	$m = 4$
1	2.24×10^7	1.96×10^7	1.95×10^7	2.25×10^7
$10^{-0.25}$	5.57×10^6	5.62×10^6	5.98×10^6	5.64×10^6
$10^{-0.5}$	1.84×10^6	1.85×10^6	1.96×10^6	1.85×10^6
$10^{-0.75}$	7.61×10^5	7.21×10^5	7.60×10^5	7.36×10^5
0.1	2.18×10^5	2.18×10^5	2.18×10^5	2.18×10^5
$10^{-1.25}$	1.01×10^5	9.93×10^4	9.93×10^4	1.01×10^5
$10^{-1.5}$	4.59×10^4	4.64×10^4	4.65×10^4	4.59×10^4
$10^{-1.75}$	2.17×10^4	2.16×10^4	2.16×10^4	2.17×10^4
0.01	7.05×10^3	7.02×10^3	7.05×10^3	7.02×10^3
$10^{-2.25}$	3.35×10^3	3.34×10^3	3.35×10^3	3.35×10^3
$10^{-2.5}$	1.11×10^3	1.11×10^3	1.11×10^3	1.11×10^3
$10^{-2.75}$	5.40×10^2	5.40×10^2	5.40×10^2	5.41×10^2
0.001	1.86×10^2	1.87×10^2	1.86×10^2	1.87×10^2

Table 6.11 Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filters in Table 6.9 where the filters are designed with fixed SFDR

β	$m = 1$	$m = 2$	$m = 3$	$m = 4$
1	3.20×10^{11}	7.74×10^{10}	4.97×10^{11}	1.07×10^{12}
$10^{-0.25}$	2.63×10^{11}	4.58×10^{10}	2.76×10^{11}	9.33×10^{11}
$10^{-0.5}$	1.10×10^{11}	2.65×10^{10}	1.52×10^{11}	7.13×10^{11}
$10^{-0.75}$	2.78×10^{11}	5.98×10^{10}	3.36×10^{11}	1.52×10^{12}
0.1	6.68×10^{10}	1.56×10^{10}	1.18×10^{11}	5.83×10^{11}
$10^{-1.25}$	1.65×10^{11}	3.86×10^{10}	2.76×10^{11}	1.25×10^{12}
$10^{-1.5}$	3.93×10^{10}	1.23×10^{10}	1.21×10^{11}	5.49×10^{11}
$10^{-1.75}$	8.36×10^{10}	2.47×10^{10}	2.34×10^{11}	1.09×10^{12}
0.01	2.92×10^{10}	1.01×10^{10}	1.16×10^{11}	1.00×10^{13}
$10^{-2.25}$	9.19×10^{10}	3.21×10^{10}	3.65×10^{11}	3.17×10^{13}
$10^{-2.5}$	3.43×10^{10}	1.47×10^{10}	1.87×10^{11}	8.89×10^{12}
$10^{-2.75}$	1.08×10^{11}	4.64×10^{10}	5.91×10^{11}	2.81×10^{13}
0.001	2.38×10^{10}	1.53×10^{10}	2.38×10^{11}	5.04×10^{12}

However, for a given set of values of N_a , N_b and η for all the values of β , the change of the condition numbers as β changes will be different. For example, if the design parameters for the FIR filters are fixed at $N_a = 47$, $N_b = 0$ and $\eta = 23$ for all the values of β , then as β decreases from 1 to 0.001, the SFDR increases from 100 dB to 158 dB and the condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ maintain about the same order of magnitude. The condition numbers for the FIR filters whose design parameters are fixed to $N_a = 47$, $N_b = 0$ and $\eta = 23$ for all the values of β are shown in Table 6.12. If the design parameters for the IIR filters are fixed at $N_a = 30$, $N_b = 4$ and $\eta = 22$ for all the values of β , then as the value of β decreases, the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ becomes more ill-conditioned. For example, when $\beta = 0.001$, the condition numbers are 1.27×10^{16} , 2.24×10^{16} , 2.26×10^{16} , and 2.63×10^{16} for $m = 1, 2, 3$, and 4 respectively and the matrices cannot be accurately inverted using Matlab software. The condition numbers for the IIR filters whose design parameters are fixed to $N_a = 30$, $N_b = 4$ and $\eta = 22$ for all the values of β are shown in Table 6.13. The design of the IIR filters is numerically less robust than that of the FIR filters. However, 100 dB SFDR is adequate for many high speed applications. For 100 dB SFDR, from our experience, the IIR filter design process is always numerically robust enough such that accurate solutions can be obtained.

Table 6.12 Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the FIR filters whose design parameters are fixed to $N_a = 47$, $N_b = 0$ and $\eta = 23$

β	$m = 1$	$m = 2$	$m = 3$	$m = 4$
1	2.24×10^7	1.96×10^7	1.95×10^7	2.25×10^7
$10^{-0.25}$	8.92×10^6	8.23×10^6	8.20×10^6	8.94×10^6
$10^{-0.5}$	6.27×10^6	5.97×10^6	5.96×10^6	6.27×10^6
$10^{-0.75}$	5.37×10^6	5.21×10^6	5.20×10^6	5.37×10^6
0.1	5.04×10^6	4.95×10^6	4.95×10^6	5.04×10^6
$10^{-1.25}$	4.92×10^6	4.87×10^6	4.87×10^6	4.92×10^6
$10^{-1.5}$	4.88×10^6	4.85×10^6	4.85×10^6	4.88×10^6
$10^{-1.75}$	4.86×10^6	4.84×10^6	4.84×10^6	4.86×10^6
0.01	4.85×10^6	4.84×10^6	4.84×10^6	4.85×10^6
$10^{-2.25}$	4.85×10^6	4.84×10^6	4.84×10^6	4.85×10^6
$10^{-2.5}$	4.85×10^6	4.84×10^6	4.84×10^6	4.85×10^6
$10^{-2.75}$	4.85×10^6	4.84×10^6	4.84×10^6	4.85×10^6
0.001	4.84×10^6	4.84×10^6	4.84×10^6	4.84×10^6

Table 6.13 Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filters whose design parameters are fixed to $N_a = 30$, $N_b = 4$ and $\eta = 22$

β	$m = 1$	$m = 2$	$m = 3$	$m = 4$
1	3.20×10^{11}	7.74×10^{10}	4.97×10^{11}	1.07×10^{12}
$10^{-0.25}$	4.41×10^{11}	1.03×10^{11}	5.76×10^{11}	1.82×10^{12}
$10^{-0.5}$	1.06×10^{12}	2.41×10^{11}	1.25×10^{12}	4.74×10^{12}
$10^{-0.75}$	2.97×10^{12}	6.66×10^{11}	3.33×10^{12}	1.39×10^{13}
0.1	8.93×10^{12}	1.98×10^{12}	9.66×10^{12}	4.25×10^{13}
$10^{-1.25}$	2.76×10^{13}	6.06×10^{12}	2.93×10^{13}	1.33×10^{14}
$10^{-1.5}$	8.66×10^{13}	1.88×10^{13}	8.97×10^{13}	4.30×10^{14}
$10^{-1.75}$	2.68×10^{14}	5.83×10^{13}	2.84×10^{14}	1.81×10^{15}
0.01	8.95×10^{14}	1.88×10^{14}	8.11×10^{14}	4.34×10^{15}
$10^{-2.25}$	2.56×10^{15}	6.06×10^{14}	4.22×10^{15}	1.52×10^{17}
$10^{-2.5}$	7.48×10^{15}	2.24×10^{15}	3.37×10^{15}	1.98×10^{16}
$10^{-2.75}$	4.90×10^{16}	2.15×10^{16}	2.61×10^{16}	1.45×10^{16}
0.001	1.27×10^{16}	2.24×10^{16}	2.26×10^{16}	2.63×10^{16}

Method-1 is used to obtain the stable IIR filters listed in Table 6.9. The numerical robustness of the IIR filters can be improved by using *Method-2*. Take $\beta = 1$ for example, another possible IIR filter design to achieve 100 dB SFDR is $N_a = 28$, $N_b = 6$, $\eta = 22$ and a uniform denominator weight of 1.789×10^{-11} . The coefficient saving is also 29.8% and the SFDR is 102.8 dB. As presented in Section 5.3, when using *Method-2*, the denominator weight matrix $\mathbf{W}_{b,m}$ which is a diagonal matrix with 1.789×10^{-11} being its diagonal elements is appended to the lower right corner of the matrix $\mathbf{W}_m$. By doing this, the weight 1.789×10^{-11} is added to the last N_b diagonal elements of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$. The condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for different values of β when the design parameters are fixed to $N_a = 28$, $N_b = 6$, $\eta = 22$ and a uniform denominator weight of 1.789×10^{-11} are listed in Table 6.14. Compared to the condition numbers for the FIR filters listed in Table 6.12, the condition numbers of the IIR filters are still 10^6 to 10^7 times larger. However, as the value of β increases, the condition numbers increase first and then maintain at about 1.16×10^{13} . The matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ can be accurately inverted for all the values of β .

Table 6.14 Condition numbers of the matrix $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$ for the IIR filters with $N_a = 28$, $N_b = 6$, $\eta = 22$ and a uniform weight of 1.789×10^{-11}

β	$m = 1$	$m = 2$	$m = 3$	$m = 4$
1	2.59×10^{12}	4.47×10^{11}	1.82×10^{12}	4.21×10^{12}
$10^{-0.25}$	2.98×10^{12}	6.50×10^{11}	2.05×10^{12}	5.44×10^{12}
$10^{-0.5}$	5.13×10^{12}	1.49×10^{12}	3.78×10^{12}	8.12×10^{12}
$10^{-0.75}$	8.01×10^{12}	3.45×10^{12}	6.69×10^{12}	1.02×10^{13}
0.1	1.01×10^{13}	6.54×10^{12}	9.45×10^{12}	1.11×10^{13}
$10^{-1.25}$	1.11×10^{13}	9.30×10^{12}	1.09×10^{13}	1.14×10^{13}
$10^{-1.5}$	1.15×10^{13}	1.07×10^{13}	1.15×10^{13}	1.15×10^{13}
$10^{-1.75}$	1.16×10^{13}	1.13×10^{13}	1.16×10^{13}	1.15×10^{13}
0.01	1.16×10^{13}	1.15×10^{13}	1.17×10^{13}	1.15×10^{13}
$10^{-2.25}$	1.16×10^{13}	1.15×10^{13}	1.17×10^{13}	1.15×10^{13}
$10^{-2.5}$	1.16×10^{13}	1.16×10^{13}	1.16×10^{13}	1.15×10^{13}
$10^{-2.75}$	1.16×10^{13}	1.16×10^{13}	1.16×10^{13}	1.15×10^{13}
0.001	1.16×10^{13}	1.15×10^{13}	1.16×10^{13}	1.16×10^{13}

6.5 Performance in the Presence of Identification Errors

The process of TIADC mismatch compensation consists of two parts, identifying the mismatch parameters and correcting for the mismatch errors. The accuracy of the identification affects the performance of the reconstruction system. In this section, we shall examine the performance of the FIR and IIR filters designed in Example 6.2 and 6.3 when the identified mismatch errors deviate from the true values, assuming the identified bandwidths and timing skews are those given in Table 6.6 and 6.8, respectively.

Let $\Omega_{c,m}$ and $\Omega_{c,m}^{true}$ denote the identified and true bandwidths, respectively, of ADC- m . Further denote the scale of the identification error as e_{iden} . Assume the true bandwidth $\Omega_{c,m}^{true}$ is related to the identified bandwidth $\Omega_{c,m}$ by

$$\Omega_{c,m}^{true} = \Omega_{c,0} + (\Omega_{c,m} - \Omega_{c,0}) \cdot RN, \quad (6.2)$$

where $\Omega_{c,0} = 2\pi f_s$ is the bandwidth of the reference ADC, and RN is a random number generated from the uniform distribution over the interval $(1 - e_{iden}, 1 + e_{iden})$. The FIR and IIR filters designed in Example 6.2 catering for the bandwidths $\Omega_{c,m}, m = 1, 2, 3, 4$ given in Table 6.6 are used to reconstruct the TIADC output when the true bandwidths are actually $\Omega_{c,m}^{true}, m = 1, 2, 3, 4$. The value of e_{iden} is varied from 0 to 0.9 at a step size of 0.05. For each value of e_{iden} , 10 thousand sets of the $\Omega_{c,m}^{true}$ values are randomly generated. The worst case SFDRs in the 10 thousand simulations for the three filters are recorded and plotted against e_{iden} in Figure 6.7. The effects of the identification error to the compensation performance of the FIR and IIR filters are almost the same. An identification error which is as small as 5 percent will degrade the SFDR by over 40 dB.

Figure 6.7 Worst case SFDR in the presence of identification error for the FIR and IIR filters designed in Example 6.2.

For the timing mismatch, we let Δr_m and Δr_m^{true} denote the identified and the true time skews, respectively, of ADC- m . Assume the true time skew Δr_m^{true} is related to the identified time skew Δr_m by the following equation

$$\Delta r_m^{true} = \Delta r_m \cdot RN. \quad (6.3)$$

In (6.3), RN is a random number generated from the uniform distribution over the interval $(1 - e_{iden}, 1 + e_{iden})$. We consider the case of $\beta = 1$. The FIR and IIR filters designed in Example 6.3 catering for the time skews $\Delta r_m, m = 1, 2, 3, 4$ given in Table 6.8 are used to reconstruct the TIADC output when the true time skews are $\Delta r_m^{true}, m = 1, 2, 3, 4$. As in the bandwidth mismatch case, the value of e_{iden} is varied from 0 to 0.9 at a step size of 0.05. For each value of e_{iden} , 10 thousand sets of the Δr_m^{true} values are randomly generated. The worst case SFDRs in the 10 thousand simulation cases for the FIR and IIR filters are recorded and plotted against e_{iden} in Figure 6.8. In the case of timing mismatch, the performances of the FIR and IIR compensation filters in the presence of identification errors are almost the same.

Figure 6.8 Worst case SFDR in the presence of identification error for the FIR and IIR filters designed in Example 6.3.

6.6 Filter Redesign When Mismatch Parameters Drift

The mismatch parameters may change due to environmental variations and component aging. The compensation filters have to be redesigned to cope with the drift in the mismatch parameters. The two methods introduced in Chapter 5 to ensure filter stability are “computer-aided” design methods, i.e.,

the group delay or the weights imposed on the denominator coefficients need to be adjusted a few times until stability is reached. Obviously it is undesirable for online design. However, it is found that, once stability is achieved, the same design parameters, i.e., the same group delay or denominator weights, for achieving stability is sufficiently robust so that stability is still guaranteed when the filters are redesigned as a result of parameter drift provided that the mismatch parameters drift is not catastrophic and the poles of the initially designed filters are not too close to the unit circle.

We shall illustrate this by using the filters designed in the previous examples. Let r_{dr} denote the drifting range of the mismatch parameter. And let RN denote a random number with a uniform distribution over the interval $(1 - r_{dr}, 1 + r_{dr})$. A new time skew value Δr_m^{new} is generated by $\Delta r_m^{new} = \Delta r_m \cdot RN$ and assigned to each sub-ADC channel (with the exception of Channel-0 which has zero time skew) and a new bandwidth value $\Omega_{c,m}^{new}$ is generated using $\Omega_{c,m}^{new} = \Omega_{c,0} + (\Omega_{c,m} - \Omega_{c,0}) \cdot RN$ and assigned to each sub-ADC channel (with the exception of Channel-0). The value of r_{dr} is varied from 0 to 0.9 with a step size of 0.05 and for each value of r_{dr} , 10 thousand sets of the new mismatch parameters are generated. Then, for each set of the new mismatch values, the FIR and IIR filters are redesigned using the same design parameters specified in the previous examples to cope with the mismatch drift. To compare the compensation performance of the redesigned filters, the same multi-tone input signal consisting of the sum of seven sinusoids, $0.0625if_s, i = 1, 2, \dots, 7$ is used. We would like to investigate how the stability and the compensation performance of the redesigned IIR filters will change as r_{dr} changes.

First consider the filters in Example 6.1 which are designed with the same total number of coefficients. The IIR filter in Figure 6.1(d) is designed with $N_a = 27, N_b = 4$ and $\eta = 20$ (stability achieved by using *Method-1*). The value of p_{max} for this filter is 0.932. Denote the set of the redesigned filters using $N_a = 27, N_b = 4$ and $\eta = 20$ as IIR-1. The IIR filter in Figure 6.1(f) is designed with $N_a = 27, N_b = 4, \eta = 15$ and a uniform denominator weight of 1.198×10^{-7} (stability achieved by using *Method-2*). The value of p_{max} for this filter is 0.930. Denote the set of the redesigned filters using the $N_a = 27, N_b = 4, \eta = 15$ and a uniform denominator weight of 1.198×10^{-7} as IIR-2. The maximum and minimum values of p_{max} in the 10 thousand simulation runs of the redesigned IIR-1

and IIR-2 filters for each value of r_{dr} are recorded and plotted in Figure 6.9. The redesigned IIR filters using the same design parameters remain stable even when the drifting range r_{dr} is increased to as high as 90%. The maximum values of p_{max} in the 10 thousand simulation runs of the IIR-1 and IIR-2 filters are 0.975 and 0.972, respectively when $r_{dr} = 0.9$. However, if the uniform denominator weight is selected to be 4.554×10^{-8} such that the initial value of p_{max} is 0.970, i.e., the IIR filter in Figure 6.1(e), the redesigned IIR filters become potentially unstable when r_{dr} is increased to 85%. The maximum value of p_{max} in the 10 thousand runs is 1.001 when $r_{dr} = 0.85$. Therefore, the design parameters must be selected such that the poles of the initially designed filter are not too close to the unit circle.

Figure 6.9 Maximum and minimum values of p_{max} of the IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the IIR filters in Fig. 6.1(d) and (f) versus r_{dr} .

The SFDRs of the compensated output using the redesigned FIR, IIR-1 and IIR-2 filters in the 10 thousand simulation runs are recorded for each value of r_{dr} . The maximum and minimum values of the SFDR in the 10 thousand runs for the three filters are plotted in Figure 6.10 as r_{dr} is increased from 0 to 90%. The variations in the mismatch parameters can be in both directions and thus the scale of the mismatch error can become either larger or smaller. As r_{dr} increases, the maximum SFDR of the redesigned filters increases and the minimum SFDR of the redesigned filters decreases. The

minimum value of the SFDR of the IIR-1 filters decreases faster than that of the FIR filters when r_{dr} is in the range from 0.5 to 0.9. Nevertheless, the compensation performances of the redesigned IIR-1 filters are still superior to that of the redesigned FIR filter.

Figure 6.10 Maximum and minimum values of SFDR of the FIR, IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the filters in Fig 6.1(c), (d) and (f) versus r_{dr} .

Example 6.2 considers bandwidth mismatch and the IIR filters are designed to achieve the same level of SFDR as the FIR filters. For each set of the new bandwidth values, FIR, IIR-1 and IIR-2 filters are redesigned using the same design parameters specified in Example 6.2. The maximum and minimum values of p_{max} of the redesigned IIR-1 and IIR-2 filters in the 10 thousand simulation runs are recorded and plotted against r_{dr} in Figure 6.11. We can see that the redesigned IIR-1 and IIR-2 filters remain stable even when r_{dr} is increased to as high as 0.9. For the redesigned IIR-1 filters, the maximum and minimum values of p_{max} vary by very small amounts when r_{dr} is increased from 0 to 0.9. For the IIR-2 filter, when r_{dr} is increased from 0 to 0.9, the maximum value of p_{max} increases from 0.950 to 0.956 and the minimum value of p_{max} drops from 0.950 to 0.788. As a result, the compensation performance of the redesigned IIR-2 filter may deteriorate when r_{dr} is large.

Next, the maximum and minimum values of the SFDR for the FIR, IIR-1 and IIR-2 filters in the 10

thousand simulations are plotted in Figure 6.12. The maximum and minimum values of the SFDR for the redesigned IIR-1 and IIR-2 filters are consistent with that for the redesigned FIR filters when r_{dr} is increased from 0 to 75%. When $r_{dr} \geq 0.80$, the minimum value of the SFDR of the IIR-2 filter is lower than that of the FIR filter. The compensation performance of the redesigned IIR-2 filter may become worse than that of the FIR filter when the drifting range is increased beyond 80%.

Figure 6.11 Maximum and minimum values of p_{max} of the IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the IIR filters in Fig. 6.4(d) and (e) versus r_{dr} .

Figure 6.12 Maximum and minimum values of SFDR of the FIR, IIR-1 and IIR-2 filters redesigned using the same design parameters as those of the filters in Fig 6.4(c), (d) and (e) versus r_{dr} .

In Example 6.3, FIR and IIR filters are designed to achieve 100 dB SFDR for a range of timing mismatch errors, i.e., for various values of β . As before, FIR and IIR-1 filters are redesigned using the same design parameters specified in Table 6.9 for each set of the new time skew values. The maximum and minimum values of p_{max} in the 10 thousand simulation runs for each value of r_{dr} and β are recorded. Figure 6.13 plots the maximum and minimum values of p_{max} against the drifting range in percentage for each value of β where $\beta = 1, 10^{-0.25}, 10^{-0.5}, \dots, 10^{-3}$. In the case where $\beta = 1$, the redesigned IIR filters become potentially unstable when the drifting range r_{dr} is increased to 75%. The maximum value of p_{max} is 1.140 when $r_{dr} = 0.75$. For all the other cases where $\beta = 10^{-0.25}, 10^{-0.5}, \dots, 10^{-3}$, the maximum (minimum) value of p_{max} increases (decreases) very slowly as r_{dr} is increased from 0 to 90% and the redesigned IIR filters remain stable for all the values of r_{dr} . The timing mismatch values for $\beta = 1$ are quite large ($\Delta r_1 = 0.36, \Delta r_2 = -0.23, \Delta r_3 = 0.27$ and $\Delta r_4 = -0.32$) and hence for the same drifting range, the absolute deviations from the initial mismatch values are also large. Therefore, the redesigned IIR-1 filters using the same design parameters are more likely to become unstable. In practical TIADC systems, the mismatches are much smaller than those in the case where $\beta = 1$. Figure 6.14 plots the maximum and minimum values of the SFDR of the FIR and IIR-1 filters in the 10 thousand simulations against r_{dr} for $\beta = 10^{-0.25}, 10^{-0.5}, \dots, 10^{-3}$. The maximum and minimum values of the SFDR of the redesigned IIR-1 filters are consistent with those of the redesigned FIR filters.

At the initial design stage of the IIR compensation filter, a few iterations are needed to determine the proper value of η or the proper denominator weights. The iterations can be performed offline. When the mismatch parameters drift and the IIR filters need to be redesigned, the same design parameters can be used for the online design of the filters. For the same compensation performance, the IIR compensation filters require about 30% less of the coefficients than the FIR filters. Hence our IIR compensation filters have lower online redesign complexities. Of course, the stability and the compensation performance of the redesigned IIR filters are of concern. Through the previous examples, we have demonstrated that the stability and the compensation performance of the redesigned IIR filters are still preserved if the mismatch drift is not catastrophic and the poles of the initially designed filters are not too close to the unit circle. It is difficult to specify a worst case bound

of the drifting range r_{dr} for which the stability and the performance of the redesigned IIR filters are guaranteed because it is affected by many factors such as the scale of the mismatch error, the method used to achieve stability, the value of p_{max} of the initially designed filter and so on. From our experience, if the design parameters are selected such that the poles of the initially designed IIR filter are not too close to the unit circle, for example, p_{max} can be constrained to be less than 0.95, the stability of the redesigned filter is guaranteed for $r_{dr} \leq 0.7$.

Figure 6.13 Maximum and minimum values of p_{max} of the IIR-1 filters redesigned using the same design parameters as those in Table. 6.9 versus r_{dr} for each value of β .

Figure 6.14 Maximum and minimum values of SFDR of the FIR and IIR-1 filters redesigned using the same design parameters as those in Table. 6.9 versus r_{dr} for each value of β .

6.7 Pipeline Implementation of the IIR Compensation Filter

Pipelined implementation of an FIR filter is straightforward. The computation of $y[n]$ can be started before the computations of the earlier output samples, $y[n-1]$, $y[n-2]$ etc. By inserting additional delays along the data path, the structure is partitioned into several pipeline stages. The computation time needed by each pipeline stage is reduced and the operating speed can thus be increased. In an IIR filter, the computation of $y[n]$ cannot be initiated, in a straightforward manner, before the computations of the previous outputs, $y[n-1]$ through $y[n-N_b]$, are completed. We shall introduce

two approaches for pipeline implementation of our proposed IIR compensation filter.

The first one is by implementing the M -PTV IIR filter in the equivalent IIR filter bank structure as shown in Figure 4.3. As demonstrated in Chapter 4, the M -PTV IIR filter operating at the rate of f_s can be represented by an equivalent MIMO system operating at the rate of f_s/M which can be further realized as a synthesis filter bank consisting of M parallel IIR filters, $R_m(z)$, $m = 0, 1, \dots, M - 1$. The IIR filter for Channel- m $R_m(z)$ is given as

$$R_m(z) = \frac{N_m(z)}{D(z^M)} \quad (6.4)$$

where $N_m(z)$ represents the numerator of the equivalent IIR filter in Channel- m and $D(z^M)$ is the common denominator of the IIR filter bank. In the equivalent filter bank representation, the output $y[n]$ is a function of $y[n - M]$, $y[n - 2M]$ etc since the common denominator is clocked at one- M th of the data rate. The IIR filter bank can then be implemented as an M -stage pipelined structure.

Take the IIR-1 filter designed in Example 6.2 for example. After the equivalent transformation, the filter coefficients of $N_m(z)$ where $m = 0, 1, \dots, 4$ and $D(z^M)$ are listed in Table. 6.15. Note that four of the filter coefficients of $N_0(z)$ (highlighted in bold case) are the same as the denominator coefficients. If $N_0(z)$ is implemented in polyphase form, the sub-filter constructed using these four coefficients together with the common denominator represents the contribution from Channel-0 input to Channel-0 output, which is one. Hence these four coefficients can be ignored in actual implementation. To implement the equivalent filter bank, a total of 144 multipliers would be needed. Compared to the initial M -PTV filter which only needs 108 multipliers, 36 more multipliers are required in exchange for higher operation speed. Nevertheless, the implementation complexity is still lower than that of the FIR compensation filter which requires 164 multiplier coefficients.

Table 6.15 Coefficient values of the equivalent filter bank for IIR-1 filter designed in Example 6.2.

$N_0(z)$	$N_1(z)$	$N_2(z)$	$N_3(z)$	$N_4(z)$	$D(z^M)$
-0.0000009361	0	0	0	0	0.6265230709
-0.0000015050	0.0000002920	0	0	0	0.0281646464
0.0000023211	-0.0000007389	0.0000001445	0	0	0.0000106600
0	0	0	0	0	
-0.0000182485	0.0000099692	-0.0000049200	0.0000021580	-0.0000008657	
0.0000678805	-0.0000398622	0.0000211360	-0.0000101306	0.0000045200	
0.0000538647	-0.0000330268	0.0000184535	-0.0000094123	0.0000045708	
-0.0000545326	0.0000344704	-0.0000200341	0.0000107913	-0.0000056235	
0	0	0	0	0	
0.0001756014	-0.0001203353	0.0000776503	-0.0000483964	0.0000306724	
-0.0005501508	0.0003840893	-0.0002540569	0.0001636840	-0.0001080590	
-0.0003956254	0.0002782712	-0.0001864925	0.0001225317	-0.0000829592	
0.0003799052	-0.0002666823	0.0001795651	-0.0001192811	0.0000820989	
0	0	0	0	0	
-0.0010921213	0.0007512710	-0.0005082941	0.0003463702	-0.0002488417	
0.0035986884	-0.0023729322	0.0015749809	-0.0010673527	0.0007702621	
0.0030627732	-0.0018421167	0.0011687233	-0.0007739123	0.0005536281	
-0.0044063772	0.0021154499	-0.0012218042	0.0007702177	-0.0005375352	
1	0	0	0	0	
-0.0147237296	1.0065513883	0.0058532004	-0.0026785844	0.0016392777	
0.0151393405	-0.0320880585	1.0149937445	0.0120112581	-0.0059044126	
0.0064019009	-0.0077378542	0.0148282325	0.9939806465	-0.0062779464	
-0.0051887967	0.0043983309	-0.0050105650	0.0103219823	0.9955253891	
0.6265230709	0	0	0	0	
-0.0066756994	0.6278276654	0.0067921609	-0.0055649278	0.0068683609	
0.0041202108	-0.0145940611	0.6301273615	0.0141633492	-0.0123439864	
0.0010519133	-0.0021058594	0.0065954223	0.6256484592	-0.0074107290	
-0.0004900531	0.0007108552	-0.0013252754	0.0046505080	0.6257651519	
0.0281646464	0	0	0	0	
-0.0001971003	0.0280381677	0.0006327808	-0.0009033458	0.0018560277	
0.0000539897	-0.0004296611	0.0279428963	0.0013503722	-0.0020276395	
0.0000053636	-0.0000272834	0.0001869906	0.0283198533	-0.0007086978	
-0.0000008327	0.0000037197	-0.0000169315	0.0001365805	0.0282652203	
0.0000106600	0	0	0	0	
-0.0000000320	0.0000103858	0.0000013156	-0.0000054049	0.0000250780	
0	-0.0000000551	0.0000100664	0.0000029072	-0.0000113629	
0	0	0.0000000223	0.0000109461	-0.0000012722	
0	0	0	0.0000000219	0.0000108291	
0	0	0	0	0	
0	0	0	0	0.0000000008	

An alternative approach for increasing the operation speed is by raising the order of the difference equation given by (3.9). This can be done by successively substituting more delayed versions of $y[n]$ into (3.9) [137]. After P times of recursion, the output $y[n]$ becomes a function of $y[n - P - 1]$, $y[n - P - 2]$ etc and the computation of $y[n]$ is required to be completed in $(P + 1)$ clock cycles

instead of one. With pipelining, the speed can be improved by $P + 1$ times. We shall demonstrate this method by using a simple example. Assume a 3-PTV IIR filter for compensating the mismatch in a 3-channel TIADC. Using the notations defined in Section 4.4, a superscript m is added to $x[n]$ and $y[n]$ to indicate that the samples $x[n]$ and $y[n]$ are contributed by Channel- m of the TIADC. The difference equations for estimating Channel-0, 1, and 2 output samples are then rewritten as follows

$$y^0[n] = a_0[0]x^0[n] + a_0[1]x^2[n-1] + b_0[1]y^2[n-1] + b_0[2]y^1[n-2], \quad (6.5)$$

$$y^1[n] = a_1[0]x^1[n] + a_1[1]x^0[n-1] + b_1[1]y^0[n-1] + b_1[2]y^2[n-2], \quad (6.6)$$

$$y^2[n] = a_2[0]x^2[n] + a_2[1]x^1[n-1] + b_2[1]y^1[n-1] + b_2[2]y^0[n-2]. \quad (6.7)$$

Delaying (6.7) by one sample and substituting $y^2[n-1]$ into (6.5), we have

$$\begin{aligned} y^0[n] = & a_0[0]x^0[n] + (a_0[1] + b_0[1]a_2[0])x^2[n-1] + b_0[1]a_2[1]x^1[n-2] \\ & + (b_0[2] + b_0[1]b_2[1])y^1[n-2] + b_0[1]b_2[2]y^0[n-3]. \end{aligned} \quad (6.8)$$

The same can be done for Channel-1 and Channel-2. We can see from (6.8) that the output $y[n]$ becomes a function of $y[n-2]$ and $y[n-3]$. The original filter consists of 2 numerator coefficients and 2 denominator coefficient for each channel. After the recursion, each channel needs 3 numerator coefficients and 2 denominator coefficients. Further substituting $y^1[n-2]$ into (6.8) results

$$\begin{aligned} y^0[n] = & a_0[0]x^0[n] + (a_0[1] + b_0[1]a_2[0])x^2[n-1] \\ & + (b_0[1]a_2[1] + a_1[0](b_0[2] + b_0[1]b_2[1]))x^1[n-2] \\ & + a_1[1](b_0[2] + b_0[1]b_2[1])x^0[n-3] \\ & + (b_0[1]b_2[2] + b_1[1](b_0[2] + b_0[1]b_2[1]))y^0[n-3] \\ & + b_1[2](b_0[2] + b_0[1]b_2[1])y^2[n-4] \end{aligned} \quad (6.9)$$

After two times of recursion, the output $y[n]$ becomes a function of $y[n-3]$ and $y[n-4]$ and the filter consists of 4 numerator coefficients and 2 denominator coefficients. The new filter can be implemented as a 3-stage pipelined structure to increase the operation speed.

6.8 Chapter Summary

In this chapter, we have demonstrated by using some design examples that: (1) IIR compensation filters with guaranteed stability can be designed by using *Method-1* and *Method-2* introduced in Chapter 5, and (2) the IIR filters achieve much better performance than the FIR counterparts when they are designed with the same total number of filter coefficients or have lower implementation complexity when they are designed to achieve the same level of SFDR. In most of the cases, at least 30% of the multipliers can be saved when using IIR filters instead of the FIR filters. Since the output samples are used in the IIR filters, the matrix to be inverted, $(\mathbf{S}_m^H \mathbf{W}_m \mathbf{S}_m)$, for the IIR filter is more ill-conditioned than the corresponding matrix in the FIR case. We have also investigated the performance degradation of the IIR compensation filters when the identified mismatch parameters deviate from their true values. It has been shown that, the performance degradation for IIR filters is almost the same as that of the FIR filters.

In real systems, mismatch parameters may drift from their initial values due to component aging, temperature variations etc. The coefficient values for different set of mismatch parameters may be pre-calculated and stored in LUTs embedded in the FPGAs and DSPs. Online filter design is another way to cope with mismatch parameter drift. Therefore, the stability and performance of the redesigned IIR filter using the same design parameters (N_a , N_b , η and denominator coefficient weight) are of concern. Fortunately, it was found that if the drifting is not catastrophic and the poles of the initially designed IIR filters are not too close to the unit circle, the stability and performance of the redesigned IIR filter using the same design parameters are still guaranteed. A suitable choice of p_{max} which leaves some margin for the poles to drift yet does not compromise much of the compensation performance would be $p_{max} = 0.95$. It is difficult to specify a given mismatch drift range for which the stability of the redesigned IIR filter is guaranteed. From our experience, if p_{max} is constrained to be less than 0.95, the stability of the redesigned filter is guaranteed for $r_{dr} \leq 0.7$.

Pipeline implementation of the M -PTV IIR filter is possible with the drawback that the implementation complexity is increased compared to the initial filter.

CHAPTER 7 : FINITE WORDLENGTH EFFECTS IN THE M -PTV IIR FILTER

7.1 Introduction

The effect of quantization has been neglected in the previous chapters for simplicity. In practice, the sampled infinite precision data has to be represented by finite number of bits resulting in a *roundoff* noise appearing at the input of the compensation filter. The process of digital filtering requires many multiplications and the wordlength of the intermediate products usually needs to be truncated to reduce the implementation complexity of the filter. Wordlength truncation is another source contributing to the total *roundoff* noise in the system. The coefficients of the compensation filter need to be quantized and the frequency response of the quantized filter is different from that of an unquantized one. This effect is called *coefficient quantization*. The third category of the quantization effect is the *limit cycle* referring to the periodic oscillations at the filter output even when there is no input. Thorough analysis of the finite wordlength effects in LTI IIR filters can be found in [138-142]. Here in this Chapter, we analysis the finite wordlength effects in the IIR mismatch compensation filter which is an LPTV system.

In Section 7.2, we establish a roundoff noise model for the proposed M -PTV IIR compensation structure. Using this model, noise propagation from the input to the output of the IIR compensation filter is investigated and compared to that of the FIR compensation filter. The analysis of the roundoff noise due to multiplier output truncation is presented in Section 7.3. The effect of coefficient quantization to the frequency response of the IIR compensation filter is studied in Section 7.4. It is

demonstrated by using an example that an IIR filter whose performance is worse than another IIR filter when both filters are implemented using infinite precision arithmetic may have a better performance when both filters are implemented using finite precision arithmetic. Section 7.5 provides an analysis of the limit cycle effect which only arises in IIR filter structures.

7.2 Roundoff Noise Introduced by Input Signal Quantization

7.2.1 Quantization Noise in an Ideal ADC

Figure 7.1 (a) An ideal ADC modelled by a sampler and a quantizer. (b) An equivalent quantization noise model for the ADC.

We start the analysis by first reviewing the quantization noise in an ideal ADC. An ADC performs two functions, samples the continuous-time signal at discrete time instants and converts the sampled signal into a digital code with finite number of bits. An ideal ADC modelled by an ideal sampler and a quantizer is shown in Figure 7.1(a). $\ddot{x}[n]$ represents the ideally sampled sequence which is of infinite precision and $\ddot{x}_Q[n]$ is the sequence obtained after quantizing $\ddot{x}[n]$ into finite number of bits. The quantization process is conveniently modelled by the addition of an error sequence $e_Q[n]$ as shown in Figure 7.1(b) and we have

$$\ddot{x}_Q[n] = \ddot{x}[n] + e_Q[n]. \quad (7.1)$$

We assume in this thesis that $\ddot{x}[n]$ is rounded to its nearest discrete value $\ddot{x}_Q[n]$. Since $\ddot{x}[n]$ is an

arbitrary real number allowed by the ADC dynamic range, the error associated with the quantizer, $e_Q[n]$, is a random variable. Further, $e_Q[n]$ is uniformly distributed over the interval $(-\Delta/2, \Delta/2)$, where Δ is the quantization step size of the ADC given by

$$\Delta = V_{FS}/(2^B - 1), \quad (7.2)$$

where V_{FS} represents the full scale voltage of the ADC and B is the number of bits of the ADC. The probability density function of $e_Q[n]$, denoted by $p(t)$ where $t = e_Q[n]$, is then

$$p(t) = \begin{cases} 1/\Delta, & -\Delta/2 \leq t \leq \Delta/2 \\ 0, & \text{elsewhere} \end{cases} \quad (7.3)$$

Therefore the quantization noise sequence, $e_Q[n]$, has zero mean and variance of

$$\sigma_e^2 = \frac{\Delta^2}{12}. \quad (7.4)$$

The variance of $e_Q[n]$, σ_e^2 , also represents the average noise power of $e_Q[n]$. Throughout this thesis, we make the assumption that the quantization noise sequence $e_Q[n]$ is a stationary white noise with zero mean and variance of $\Delta^2/12$. We further assume that it is uncorrelated with the input signal $\tilde{x}[n]$.

7.2.2 Quantization Noise in a Non-ideal TIADC

This section considers the quantization noise introduced by the TIADC and its propagation from the input to the output of the M -PTV IIR compensation filter. In the TIADC structure shown in Figure 7.2(a), the noise sources associated with the quantizers in the M sub-ADCs are modelled by the addition of the noise sequences $e_{x,m}[k]$ where $m = 0, 1, \dots, M-1$. Assume that the noise sequence $e_{x,m}[k]$ is uncorrelated to the unquantized sub-sequence $x_m[k]$ for $m = 0, 1, \dots, M-1$. We further assume that the M noise sources are stationary white noises with zero mean and uncorrelated to each other. The M noise sources also have the same variance because the M sub-ADCs have the same full scale voltage and resolution. Denote this noise variance as $\sigma_{e_x}^2$ and it is given by (7.4). When the M

quantized sub-sequences are interleaved by the digital multiplexer, the error signals $e_{x,m}[k], m = 0, 1, \dots, M - 1$ forms into one stationary white noise with zero mean and variance $\Delta^2/12$. Denote the interleaved error sequence as $e_x[n]$ and the quantized TIADC output sequence as $x_Q[n]$. We have

$$x_Q[n] = x[n] + e_x[n], \quad (7.5)$$

where $x[n]$ is the unquantized TIADC output sequence and it is uncorrelated to $e_x[n]$.

Figure 7.2 (a) Quantization noise model for a TIADC. (b) Quantization noise $e_x[n]$ propagates through the M -PTV IIR compensation filter $H_n(z)$ and appears at the output as $e_y[n]$.

Our compensation system is an M -PTV IIR filter whose z -transform at time instant n is denoted by $H_n(z)$. The quantization noise $e_x[n]$ propagates through $H_n(z)$ and appears at the output as a different error sequence $e_y[n]$ as shown in Figure 7.2(b). To analyze the noise propagation in the proposed compensation structure, we start by reviewing the noise propagation in an LTI system [138-140]. Let $e_Q[n]$ be a white noise with zero mean and variance σ_e^2 . When $e_Q[n]$ propagates through an LTI system with transfer function $F(z)$, the signal appearing at the output of the LTI system is also a zero mean white noise. Let $\sigma_{e_F}^2$ denote the variance of the output noise of $F(z)$ in response to the input noise $e_Q[n]$ and

$$\sigma_{e_F}^2 = \sigma_e^2 \sum_{n=-\infty}^{\infty} |f[n]|^2, \quad (7.6)$$

where $f[n]$ is the impulse response of $F(z)$ and the summation term represents the energy of $F(z)$. After propagating through an LTI system, the variance of the noise is amplified by the energy of the system. Using discrete-time Fourier transform (DTFT), (7.6) is equivalently written as

$$\begin{aligned} \sigma_{e_F}^2 &= \sigma_e^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\omega})|^2 d\omega \\ &= \sigma_e^2 \|F(e^{j\omega})\|_2^2, \end{aligned} \quad (7.7)$$

where $F(e^{j\omega})$ is obtained by evaluating $F(z)$ at $z = j\omega$ and $\|F(e^{j\omega})\|_2$ denotes the L_2 norm of $F(e^{j\omega})$.

Figure 7.3 Quantization noise model for the equivalent MIMO system of the M -PTV IIR filter.

As demonstrated in Chapter 4, an M -PTV IIR filter can alternatively be interpreted as an M -input, M -output MIMO system. The MIMO system itself is LTI and operating at the rate of f_s/M . We shall make use of the equivalent MIMO representation to analyze the noise propagation in our IIR compensation filter. As shown in Figure 7.3, the input noise vector to the equivalent MIMO system is

$$\mathbf{e}_x[k] = [e_{x,0}[k], e_{x,1}[k], \dots, e_{x,M-1}[k]]^T. \quad (7.8)$$

Let $e_{y,j}[k]$ be the noise at the j -th output port caused by the input noise vector propagating through the MIMO system. The output noise vector, $\mathbf{e}_y[k]$, is then formed as

$$\mathbf{e}_y[k] = [e_{y,0}[k], e_{y,1}[k], \dots, e_{y,M-1}[k]]^T. \quad (7.9)$$

Let $\mathbf{E}_x(z)$ and $\mathbf{E}_y(z)$ denote the z -transforms of $\mathbf{e}_x[k]$ and $\mathbf{e}_y[k]$, respectively, then we have

$$\mathbf{E}_x(z) = [E_{x,0}(z), E_{x,1}(z), \dots, E_{x,M-1}(z)]^T, \quad (7.10)$$

$$\mathbf{E}_y(z) = [E_{y,0}(z), E_{y,1}(z), \dots, E_{y,M-1}(z)]^T, \quad (7.11)$$

where $E_{x,m}(z) = \sum_{k=-\infty}^{\infty} e_{x,m}[k]z^{-k}$ and $E_{y,m}(z) = \sum_{k=-\infty}^{\infty} e_{y,m}[k]z^{-k}$ are the z -transforms of $e_{x,m}[k]$ and $e_{y,m}[k]$, respectively. The z -transforms of the output noise vector and the input noise vector are then related by

$$\mathbf{E}_y(z) = \mathbf{H}(z)\mathbf{E}_x(z), \quad (7.12)$$

where $\mathbf{H}(z)$ is the transfer function matrix of the equivalent MIMO system and is given by (4.24).

From (7.12), we can write the following

$$E_{y,j}(z) = \sum_{i=0}^{M-1} H_{j,i}(z)E_{x,i}(z), \quad (7.13)$$

where $H_{j,i}(z)$ is the entry of $\mathbf{H}(z)$ at the j -th row and i -th column. Since the transfer function matrix $\mathbf{H}(z)$ is LTI and the M input noise sources $e_{x,i}[k], i = 0, 1, \dots, M-1$ are white noises with zero mean and variance $\sigma_{e_x}^2$, the noise $e_{y,j}[k]$ at the j -th output is also an white noise with zero mean. However, the noise variances at different output ports are in general different. Let $\sigma_{e_y,j}^2$ be the variance of $e_{y,j}[k]$.

It can be computed by

$$\sigma_{e_y,j}^2 = \sigma_{e_x}^2 \sum_{i=0}^{M-1} \|H_{j,i}(e^{j\omega})\|_2^2, \quad (7.14)$$

where $H_{j,i}(e^{j\omega})$ is $H_{j,i}(z)$ being evaluated at $z = j\omega$ and $\|H_{j,i}(e^{j\omega})\|_2$ denotes the L_2 norm of $H_{j,i}(e^{j\omega})$. The variance at the j -th output is the input variance being amplified by the sum of the energy of the elements in the j -th row of the transfer function matrix $\mathbf{H}(z)$. The interleaved output error sequence $e_y[n]$ is in general not stationary since its variance is time-varying.

Example 7.1

We shall verify the above analysis of noise propagation in the IIR compensation filter using the IIR filters designed in Example 6.1. Noise propagation analysis in the FIR compensation filter is also presented for comparison. Denote $g_{noise,j}$ as the sum of the energy of the elements at the j -th row of the transfer function matrix $\mathbf{H}(z)$, i.e., $g_{noise,j} = \sum_{i=0}^{M-1} \|H_{j,i}(e^{j\omega})\|_2^2$. $g_{noise,j}$ represents the noise gain in the j -th channel of the M -PTV IIR compensation filter. The values of $g_{noise,j}$ where $j = 0, 1, \dots, M-1$ are computed from the filter coefficients of the FIR and IIR filters given in Tables 6.2-4.5 and are tabulated in Table 7.1.

Table 7.1 Noise amplifications at the output ports of the filters designed in Example 6.1. The values are obtained through analysis.

	$g_{noise,0}$	$g_{noise,1}$	$g_{noise,2}$	$g_{noise,3}$	$g_{noise,4}$
FIR	1.000	1.301	1.472	1.711	2.031
IIR-1 ($p_{max} = 0.932$)	1.000	1.427	1.683	1.802	1.954
IIR-2 ($p_{max} = 0.970$)	1.000	1.559	1.976	2.231	2.163
IIR-2 ($p_{max} = 0.930$)	1.000	1.461	1.727	1.835	1.945

To verify the results in Table 7.1, the output noise variances of each filter when the input is excited with a white noise are examined. Assume a 10-bit quantization, then the input noise has zero mean and variance of $\sigma_{e_x}^2 = 2^{-20}/12 = 7.947 \times 10^{-8}$. To simulate the input noise, the filters are excited by a long stream of pseudo random numbers drawn from the uniform distribution over the interval $(-2^{-11}, 2^{-11})$. The noise variances at the five output ports, $\sigma_{e_y,m}^2$, $m = 0, 1, \dots, M-1$, for the FIR and

the IIR compensation filters are listed in Table 7.2 and the noise amplification in each channel, $\sigma_{e_y,m}^2/\sigma_{e_x}^2$, $m = 0, 1, \dots, M-1$ are shown in Table 7.3.

Table 7.2 Noise variances at the output ports of the filters designed in Example 6.1. Assume 10-bit quantization noise input.

	$\sigma_{e_y,0}^2 \times 10^8$	$\sigma_{e_y,1}^2 \times 10^8$	$\sigma_{e_y,2}^2 \times 10^8$	$\sigma_{e_y,3}^2 \times 10^8$	$\sigma_{e_y,4}^2 \times 10^8$
FIR	7.966	10.286	11.735	13.581	16.193
IIR-1 ($p_{max} = 0.932$)	7.962	11.284	13.505	14.220	15.803
IIR-2 ($p_{max} = 0.970$)	7.989	12.477	15.693	17.736	17.351
IIR-2 ($p_{max} = 0.930$)	7.870	11.614	13.737	14.564	15.562

Table 7.3 Noise amplifications at the output ports of the filters designed in Example 6.1. The values are obtained through simulations.

	$\sigma_{e_y,0}^2/\sigma_{e_x}^2$	$\sigma_{e_y,1}^2/\sigma_{e_x}^2$	$\sigma_{e_y,2}^2/\sigma_{e_x}^2$	$\sigma_{e_y,3}^2/\sigma_{e_x}^2$	$\sigma_{e_y,4}^2/\sigma_{e_x}^2$
FIR	1.000	1.291	1.473	1.705	2.033
IIR-1 ($p_{max} = 0.932$)	1.002	1.421	1.700	1.790	1.990
IIR-2 ($p_{max} = 0.970$)	1.005	1.570	1.975	2.232	2.184
IIR-2 ($p_{max} = 0.930$)	0.992	1.463	1.731	1.835	1.961

The noise amplifications listed in Table 7.3 which are obtained by injecting a white noise at the input and calculating the noise variances at the output ports agree with those in Table 7.1 which are obtained through analysis. For an FIR compensation filter, the output noise power should be about the same as the input noise power for reasonably small mismatches. In Table 7.1 and 7.3, the noise amplifications in Channel-1, 2, 3, and 4 of the FIR filter are greater than one and this is attributed to the large mismatch parameters simulated in Example 6.1. The output amplification by the IIR-1 filter is only slightly larger than that by the FIR filter. The noise amplification by the IIR-2 filter with $p_{max} = 0.970$ is greater than that by the IIR-2 filter with $p_{max} = 0.930$. The larger noise amplification may be due to the larger pole radius. The output noise power of an IIR filter increases as the filter poles get closer to the unit circle. To verify the relationship between the pole radius and the noise amplification of our IIR compensation filter, the denominator coefficient weights of the IIR-2

filter in Example 6.1 are controlled carefully such that the value of p_{max} of the redesigned filter is gradually increased from 0.90 to 0.99 with a step size of 0.01. Then the noise amplification in all the channels except the reference channel, $g_{noise,j}, i = 1, 2, 3, 4$ are plotted in Figure 7.4 against the maximum pole radius p_{max} . The noise amplification in the reference channel always equals to one. As expected, the noise amplification $g_{noise,j}$ in general increases with the value of p_{max} . As p_{max} approaches one, $g_{noise,j}$ increases rapidly. Therefore, from the point of view of the noise propagation, it is desirable that the poles of the designed IIR filter are not too close to the unit circle.

Figure 7.4 Noise amplification in each channel of the IIR compensation filter with respect to the maximum pole radius p_{max} .

For reasonably small mismatch values, the noise amplifications by both the FIR and IIR filters are close to unity. We shall demonstrate this using the filters designed in Example 6.2. Using the equivalent MIMO representation, the noise amplification in each channel, $g_{noise,j}$ where $j = 0, 1, \dots, 4$ are computed from the transfer function matrix $\mathbf{H}(z)$ and are tabulated in Table 7.4. The values of $g_{noise,m}$ where $m = 0, 1, \dots, 4$ are very close to unity for the FIR, IIR-1 and IIR-2 filters when the mismatch errors are small. We also assume a 10-bit quantization noise with variance of 7.947×10^{-8} is injected at the input of the filters and calculate the noise power at each output port of the three

filters. The noise gains obtained through simulation, $\sigma_{e_y,m}^2/\sigma_{e_x}^2$ where $m = 0,1,\dots,4$, are listed in Table 7.5. The noise amplifications of the compensation filters obtained through simulation agree with those obtained through analysis.

Table 7.4 Noise amplifications at the output ports of the filters designed in Example 6.2. The values are obtained through analysis.

	$\mathcal{I}_{noise,0}$	$\mathcal{I}_{noise,1}$	$\mathcal{I}_{noise,2}$	$\mathcal{I}_{noise,3}$	$\mathcal{I}_{noise,4}$
FIR	1.000	1.011	1.026	0.991	0.993
IIR-1	1.000	1.012	1.030	0.989	0.992
IIR-2	1.000	1.014	1.032	0.988	0.991

Table 7.5 Noise amplifications at the output port of the filters designed in Example 6.2. The values are obtained through simulations.

	$\sigma_{e_y,0}^2/\sigma_{e_x}^2$	$\sigma_{e_y,1}^2/\sigma_{e_x}^2$	$\sigma_{e_y,2}^2/\sigma_{e_x}^2$	$\sigma_{e_y,3}^2/\sigma_{e_x}^2$	$\sigma_{e_y,4}^2/\sigma_{e_x}^2$
FIR	1.007	1.013	1.022	0.985	0.994
IIR-1	1.000	1.012	1.023	0.995	0.992
IIR-2	1.007	1.013	1.023	0.984	0.999

7.3 Roundoff Noise Introduced by Multiplier Output Truncation

In this section, we shall investigate the roundoff noise introduced by multiplier output truncation. Consider the Direct Form I implementation of the M -PTV IIR compensation filter shown in Figure 7.5. For fixed-point arithmetic, the wordlength of the product of two numbers, one with wordlength a and the other with wordlength b , is $a + b$. The product of the signal and the coefficients needs to be truncated to reduce the filter implementation complexity. Let $e_{a,p}[n], p = 0,1,\dots,N_a - 1$ and $e_{b,q}[n], q = 1,2,\dots,N_b$ denote the errors caused by wordlength truncation of the numerator multiplier outputs and denominator multiplier outputs, respectively. The noise model is shown in Figure 7.5(a). Assume the outputs of all the multipliers are truncated to B_h bits. Then $e_{a,p}[n]$ and $e_{b,q}[n]$ are white

noises with zero mean and variances σ_h^2 where

$$\sigma_h^2 = \frac{1}{12(2^{B_h} - 1)^2}. \quad (7.15)$$

Figure 7.5 Noise model for multiplier output truncation in an M -PTV IIR filter. (a) The errors due to multiplier output truncation are modelled by the addition of the noise sources $e_{a,p}[n]$, $p=0, 1, \dots, N_a-1$ and $e_{b,q}[n]$, $q=0, 1, \dots, N_b$. (b) The multiplier output truncation noise sources are replaced by a single noise source $e_c[n]$.

Further assume that $e_{a,p}[n], p = 0, 1, \dots, N_a - 1$ and $e_{b,q}[n], q = 1, 2, \dots, N_b$ are uncorrelated and each of them is in turn uncorrelated with the input signal. With simple manipulations, the N noise sources in Figure 7.5(a) can be replaced by a single noise source $e_c[n]$ as shown in Figure 7.5(b) where $N = N_a + N_b$. We have

$$e_c[n] = \sum_{p=0}^{N_a-1} e_{a,p}[n] + \sum_{q=1}^{N_b} e_{b,q}[n]. \quad (7.16)$$

When designing the TIADC mismatch compensation filters, it is common to choose Channel-0 as the reference and the filters are designed to compensate for the relative mismatches in the rest of the channels with respect to that in Channel-0. Therefore, for Channel-0, the numerator coefficient $a_0[p]$ is one when $p = \eta$ and zero otherwise; the denominator coefficients values are all zero. Usually $B_h \geq B$, where B is the number of bits of the sub-ADCs, and hence the roundoff noise caused by multiplier output truncation in Channel-0 is 0, i.e., $e_{a,p}[n] = 0$ and $e_{b,q}[n] = 0$ for $\langle n \rangle M = 0$ where $\langle n \rangle M$ denote the n modulo M operation. Further denote the noise due to multiplier output truncation in Channel- i as $e_{c,i}[k]$. Then $e_c[n]$ can be view as the sequence which is obtained by multiplexing the M sequences, $e_{c,i}[k], i = 0, 1, \dots, M - 1$. $e_{c,i}[k]$ has zero mean and variance $\sigma_{c,i}^2$ which is given by

$$\sigma_{c,i}^2 = \begin{cases} 0, & i = 0 \\ N\sigma_h^2, & i = 1, 2, \dots, M - 1 \end{cases} \quad (7.17)$$

The output noise due to multiplier output truncation is obtained by passing $e_c[n]$ through a new M -PTV IIR filter whose denominator coefficients are the same as those of the IIR compensation filter and numerator coefficient is simply one. The new M -PTV IIR filter is shown inside the dashed box in Figure 7.5(b). Denote the transfer function of the new M -PTV IIR filter at time instant n as $H_n^{deno}(z)$, then

$$H_n^{deno}(z) = \frac{1}{1 + \sum_{q=1}^{N_b} b_n[q]z^{-q}}. \quad (7.18)$$

The noise signal $e_c[n]$ propagating through the M -PTV system $H_n^{deno}(z)$ can be modelled by the following noise vector propagating through the equivalent LTI MIMO system of $H_n^{deno}(z)$.

$$\mathbf{e}_c[k] = [e_{c,0}[k], e_{c,1}[k], \dots, e_{c,M-1}[k]]^T \quad (7.19)$$

Let $\mathbf{H}^{deno}(z)$ denote the transfer function matrix of the equivalent MIMO system of $H_n^{deno}(z)$ and it can be derived by following the same procedures demonstrated in Chapter 4. The output noise at the j -th output of the MIMO system $\mathbf{H}^{deno}(z)$ is a white noise with zero mean and variance $\sigma_{cy,j}^2$ which is given by

$$\begin{aligned} \sigma_{cy,j}^2 &= \sum_{i=0}^{M-1} \sigma_{c,i}^2 \|H_{j,i}^{deno}(e^{j\omega})\|_2^2 \\ &= N\sigma_h^2 \sum_{i=1}^{M-1} \|H_{j,i}^{deno}(e^{j\omega})\|_2^2 \end{aligned} \quad (7.20)$$

where $\|H_{j,i}^{deno}(e^{j\omega})\|_2$ is the L_2 norm of $H_{j,i}^{deno}(e^{j\omega})$ and $H_{j,i}^{deno}(e^{j\omega})$ is the entry of $\mathbf{H}^{deno}(z)$ at the j -th row and the i -th column being evaluated at $z = j\omega$. When the noise signals at the M output ports are interleaved, the variance of the resulting noise signal is time-varying and periodic with period M .

Example 7.2

We shall verify the above analysis of the roundoff noise due to multiplier output wordlength truncation by using an example. Consider the same 5-channel TIADC and the FIR and IIR compensation filters in Example 6.1. For ease of illustration, construct another $M \times M$ matrix, $\|\mathbf{H}^{deno}(e^{j\omega})\|^2$, whose element at the j -th row and i -th column is $\|H_{j,i}^{deno}(e^{j\omega})\|_2^2$. The matrices $\|\mathbf{H}^{deno}(e^{j\omega})\|^2$ for the three IIR filters in Example 6.1 are computed from the denominator

coefficients and shown in Table 7.6. For the FIR filter, the denominator coefficients are zero and the energy matrix $\|\mathbf{H}^{deno}(e^{j\omega})\|^2$ is simply an $M \times M$ identity matrix. Hence the roundoff noise due to multiplier output wordlength truncation is not amplified by the FIR filter. On the contrary, for the IIR filters, the energies of the elements of $\mathbf{H}^{deno}(z)$ can be much greater than unity except those in the first row and first column. Therefore the noise powers in all the channels except Channel-0 are amplified by a significant factor. The element at the j -th row and i -th column of $\mathbf{H}^{deno}(z)$ represents the contribution from Channel- i output to Channel- j output. Channel-0 has all zero denominator coefficients meaning that the contributions from other channels' output to Channel-0 output are zero. Hence $H_{0,i}^{deno}(z) = 0$ where $i = 1, 2, \dots, 4$. The denominator coefficients of Channels 1 to 4 corresponding to the contribution from Channel-0 outputs are zeros as well (the reason has been discussed in Section 3.4). Hence $H_{j,0}^{deno}(z) = 0$ where $j = 1, 2, \dots, 4$.

Table 7.6 The energy of each element of the transfer function matrix $\mathbf{H}^{deno}(z)$ for the IIR filters designed in Example 6.1.

	$\ \mathbf{H}^{deno}(e^{j\omega})\ ^2$
IIR-1 (Table 6.3)	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 5.408 & 4.551 & 3.402 & 5.109 \\ 0 & 16.748 & 13.590 & 9.726 & 15.120 \\ 0 & 30.358 & 26.310 & 17.361 & 26.157 \\ 0 & 19.054 & 17.710 & 11.928 & 16.921 \end{bmatrix}$
IIR-2 (Table 6.4)	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3.775 & 1.566 & 1.999 & 2.522 \\ 0 & 8.394 & 5.387 & 5.068 & 7.870 \\ 0 & 16.880 & 9.432 & 10.476 & 14.004 \\ 0 & 9.865 & 6.118 & 6.379 & 9.669 \end{bmatrix}$
IIR-2 (Table 6.5)	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1.614 & 0.369 & 0.519 & 0.593 \\ 0 & 2.128 & 1.955 & 0.987 & 2.002 \\ 0 & 4.581 & 2.851 & 2.785 & 3.153 \\ 0 & 2.362 & 1.807 & 1.465 & 2.839 \end{bmatrix}$

Assume the outputs of all the numerator and denominator multipliers are truncated to 14 bits, then $\sigma_h^2 = 3.104 \times 10^{-10}$. The noise variance due to multiplier output truncation at each output port of the FIR and IIR filters are calculated using (7.20) and Table 7.6 and are tabulated in Table 7.7. Note that the noise variances at Channels 1 to 4 outputs for the FIR filter are simply $N\sigma_h^2 = 9.313 \times 10^{-9}$. To

simulate the effect of multiplier wordlength truncation, the input signal is set to zero; then 30 streams of pseudo random numbers drawn from the uniform distribution over the interval $[-2^{-15}, 2^{-15}]$ are added to the 30 multiplier outputs. The output noise variance in each channel is recorded and shown in Table 7.8. The noise variances in Table 7.8 which are obtained through simulation agree with the results obtained through analysis in Table 7.7. It is observed from both Table 7.7 and 7.8 that the output noise variances in Channels 1 to 4 for the IIR filters can be several tens of times of that for the FIR filter.

To reduce the effect caused by multiplier wordlength truncation, the wordlength after truncation, B_h , can be set to be several bits longer than the input quantization bits so that the total roundoff noise at the output is not amplified by a big factor. Take the IIR-1 filter in Example 6.1 for example, if the multiplier output wordlength is set to 18 bits, then the noise variances at the five output ports are 0, 6.719×10^{-10} , 2.008×10^{-9} , 3.645×10^{-9} , and 2.387×10^{-9} . Compared to the output noise due to input quantization listed in Table 7.2, the noise contributed by multiplier output truncation is very small and can be ignored.

Table 7.7 The noise variances due to multiplier output wordlength truncation for the FIR and IIR filters designed in Example 6.1. The results are obtained through analysis.

	$\sigma_{cy,0}^2$	$\sigma_{cy,1}^2$	$\sigma_{cy,2}^2$	$\sigma_{cy,3}^2$	$\sigma_{cy,4}^2$
FIR	0	9.313×10^{-9}	9.313×10^{-9}	9.313×10^{-9}	9.313×10^{-9}
IIR-1 ($p_{max} = 0.932$)	0	1.720×10^{-7}	5.139×10^{-7}	9.331×10^{-7}	6.111×10^{-7}
IIR-2 ($p_{max} = 0.970$)	0	9.185×10^{-8}	2.488×10^{-7}	4.730×10^{-7}	2.983×10^{-7}
IIR-2 ($p_{max} = 0.930$)	0	2.883×10^{-8}	6.585×10^{-8}	1.245×10^{-7}	7.890×10^{-8}

Table 7.8 The noise variances due to multiplier output wordlength truncation for the FIR and IIR filters designed in Example 6.1. The results are obtained through simulation.

	$\sigma_{cy,0}^2$	$\sigma_{cy,1}^2$	$\sigma_{cy,2}^2$	$\sigma_{cy,3}^2$	$\sigma_{cy,4}^2$
FIR	0	9.233×10^{-9}	9.310×10^{-9}	9.362×10^{-9}	9.219×10^{-9}
IIR-1 ($p_{max} = 0.932$)	0	1.721×10^{-7}	5.142×10^{-7}	9.341×10^{-7}	6.124×10^{-7}
IIR-2 ($p_{max} = 0.970$)	0	9.091×10^{-8}	2.470×10^{-7}	4.690×10^{-7}	2.958×10^{-7}
IIR-2 ($p_{max} = 0.930$)	0	2.870×10^{-8}	6.556×10^{-8}	1.233×10^{-7}	7.877×10^{-8}

In Example 6.2, the FIR and IIR filters are designed to achieve the same compensation performance. For compensating one channel output, the FIR filter requires 41 multipliers whereas the IIR-1 filter only requires 27 multipliers. 34% of the multipliers are saved when the IIR-1 filter is used instead of the FIR filter. Since the multiplier output truncation noise is amplified by the feedback part of the IIR compensation filter, the IIR filter would need longer multiplier output truncation wordlength in order to achieve the same noise power as that in the FIR filter. For example, assume the multiplier outputs in the FIR filter implementation are truncated to 14 bits, and then the output noise variance is 1.273×10^{-8} for Channel 1 to 4. For the IIR-1 filter to have about the same amount of output noise, the multiplier outputs need to be truncated to 17 bits wordlength, and then the output noise variances at Channel 1, 2, 3 and 4 are 2.670×10^{-9} , 1.141×10^{-8} , 2.765×10^{-9} and 1.409×10^{-9} , respectively. The implementation complexity may be increased when using the IIR-1 filter.

In actual implementation, multiplications are the major resource consuming operations in terms of the time, area and power consumption. Also, for online filter design, the amount of the computations needed is determined by the filter order. For the design of both the FIR and IIR compensation filters, M separate matrix inversions need to be calculated where each matrix is with size $N \times N$ and N is the number of the filter coefficients to be optimized. The amount of multiplications needed to compute the matrix inverse is $\mathcal{O}(N^3)$ if Gauss-Jordan elimination is used. Hence, the increase in the implementation complexity due to the longer truncation wordlength is very small when compared to the amount of the resources being saved by the 34% less number of multipliers.

7.4 Coefficient Quantization

When the coefficients of the compensation filter are quantized into finite wordlength, the compensation performance of the filter will deteriorate. If the designed IIR compensation filter is stable yet with poles close to the unit circle, it may become unstable after coefficient quantization. Under finite wordlength condition, the coefficient sensitivity of an IIR filter depends on its pole radius. In general, the nearer is the pole to the unit circle, the higher is the coefficient sensitivity. We

introduced two methods, namely, *Method-1* and *Method-2*, in Chapter 5 to constrain the pole radius of the IIR filter. Constraining the pole radius beyond that is required to achieve stability will degrade the filter's performance when implemented using infinite precision arithmetic. Nevertheless, constraining the pole radius beyond that is required to achieve stability will produce filters whose poles are further away from the unit circle; the filters are more robust under finite wordlength condition and their performances degrade slower as the wordlength decreases. Thus, an IIR filter whose performance is worse than another IIR filter when both filters are implemented using infinite precision arithmetic may have a better performance when both filters are implemented using finite precision arithmetic. In this section, we shall use some examples to illustrate this fact.

Example 7.3

Consider the IIR filters designed in Example 6.1. For the IIR filter whose stability is achieved by using *Method-1*, the larger is the value of η , the smaller is the pole radius and the more robust is the filter under finite wordlength condition. Figure 7.6 shows the performances of two IIR compensation filters. One of the IIR filter is designed with $\eta = 20$ and the other is with $\eta = 21$. Here, we define a new measurement for the compensation performance, $\sigma_{spurious}^2$, as the sum of squares of the magnitudes of the spurious components. Figure 7.6(a) corresponds to the design with $\eta = 20$ and implemented using infinite precision arithmetic. Figure 7.6(a) is the same as Figure 6.1(d); it is reproduced here to facilitate comparison. The value of p_{max} for the IIR filter in Figure 7.6(a) is 0.932 and the value of $\sigma_{spurious}^2$ is -87.5 dB. Figure 7.6(b) corresponds to the design with $\eta = 21$ and implemented using infinite precision arithmetic. The value of p_{max} for the IIR filter in Figure 7.6(b) is 0.880 and the value of $\sigma_{spurious}^2$ is -83.2 dB. It can be seen that the performance of the design with $\eta = 21$ is worse than that with $\eta = 20$ when the filters are implemented using infinite precision arithmetic. However, the filter with $\eta = 21$ has a smaller pole radius and so is expected to be more robust under finite wordlength arithmetic. Figure 7.6(c) and (d) correspond to the design with $\eta = 20$ and $\eta = 21$, respectively, when the coefficient wordlengths are reduced to 18-bit. The values of $\sigma_{spurious}^2$ are -64.8 dB and -70.1 dB, respectively, for the design with $\eta = 20$ and $\eta = 21$. It can be

seen that the performance of the design with $\eta = 21$ is better than that with $\eta = 20$ when implemented using 18-bit coefficient wordlength.

Figure 7.6 Simulated output frequency spectrum of the IIR filters. Stability is achieved using *Method-1*. (a) Infinite precision and $\eta = 20$. (b) Infinite precision and $\eta = 21$. (c) 18-bit coefficient wordlength and $\eta = 20$. (d) 18-bit coefficient wordlength and $\eta = 21$.

Figure 7.7 plots the values of $\sigma_{spurious}^2$ of the FIR filter and the two IIR filters with $\eta = 20$ and $\eta = 21$ as the number of coefficient quantization bits is increased from 5 to 30. The values of $\sigma_{spurious}^2$ for the three filters decrease with increasing number of quantization bits before they reach to their respective minimum values which are the $\sigma_{spurious}^2$ values achieved when infinite precision arithmetic is used. The number of quantization bits needed to reach about the same performance as that under infinite precision arithmetic are 16-bit, 23-bit and 22-bit, respectively, for the FIR filter, IIR

filter with $\eta = 20$ and IIR filter with $\eta = 21$. The FIR filter has better performance than that of the two IIR filters if the number of quantization bits is less than or equal to 15. Although the IIR filter with $\eta = 20$ has better performance when implemented using infinite precision, its performance is worse than the IIR filter with $\eta = 21$ if their coefficients are implemented with less than 18-bit precision. The IIR filter with smaller pole radius is more robust under finite wordlength condition.

Figure 7.7 $\sigma_{spurious}^2$ of the FIR filter and the IIR filters with $\eta = 20$ and $\eta = 21$ versus coefficient wordlength.

For the IIR filter whose stability is achieved by using *Method-2*, imposing greater weights on the denominator coefficients would result in filters with smaller pole radius. The IIR filter in 6.1(e) has $\sigma_{spurious}^2$ of -77.7 dB and p_{max} of 0.970. The spectrum of the compensated output is reproduced in Figure 7.8(a) to facilitate comparison. Another IIR filter with smaller pole radius is obtained by imposing a uniform weight of 3.058×10^{-7} on all the denominator coefficients. The value of p_{max} is 0.900 and the value of $\sigma_{spurious}^2$ is -70.6 dB. The spectrum of the compensated output is shown in Figure 7.8(b). It can be seen that the performance of the design with a smaller pole radius is worse than the design with a larger pole radius when the filters are implemented using infinite precision arithmetic. When the coefficient wordlengths are reduced to 16-bit, the performance of the design with p_{max} equal to 0.970 and 0.900 are shown in Figure 7.8(c) and Figure 7.8(d), respectively. The

values of $\sigma_{spurious}^2$ are -55.3 dB and -63.8 dB, respectively for the design with $p_{max} = 0.970$ and $p_{max} = 0.900$. The performance of the design with smaller value of p_{max} is better when the coefficient wordlengths of both filters are reduced to 16-bit.

Figure 7.8 Simulated output frequency spectrum of the IIR filters. Stability is achieved using *Method-2*. (a) Infinite precision with $p_{max} = 0.970$. (b) Infinite precision with $p_{max} = 0.900$. (c) 16-bit coefficient wordlength with $p_{max} = 0.970$. (d) 16-bit coefficient wordlength with $p_{max} = 0.900$.

Figure 7.9 plots the values of $\sigma_{spurious}^2$ of the FIR filter and the two IIR filters as the number of coefficient quantization bits is increased from 5 to 30. The values of $\sigma_{spurious}^2$ of the three filters decrease with increased coefficient wordlength until they reach to their respective minimums, i.e., the $\sigma_{spurious}^2$ values achieved under infinite precision arithmetic. The FIR filter has better performance than that of the two IIR filters if the number of quantization bits is less than 15. The IIR filter with

smaller pole radius, $p_{max} = 0.900$, is more robust under finite wordlength condition and its performance is better than that of the IIR filter with $p_{max} = 0.970$ when the coefficients are implemented with less than 17-bit precision.

Figure 7.9 $\sigma_{spurious}^2$ of the FIR filter and the IIR filters with $p_{max} = 0.970$ and $p_{max} = 0.900$ versus coefficient wordlength.

7.5 Limit Cycle

As demonstrated in previous sections, signal quantization, including input signal quantization and multiplier output truncation, in the IIR compensation filter generates a random noise at the filter output. Signal quantization in a recursive filter may also cause periodic oscillations called limit cycle oscillations. The oscillations are self sustained when there is no input.

Consider the implementation of a first order recursive filter $y[n] = x[n] - by[n - 1]$ shown in Figure 7.10. Assume that the recursive filter is stable with $b = 0.7$. Suppose that $x[n]$ is an impulse and there is no quantizer, then the output $y[n]$ will decay to zero eventually. However, the result is different if a

quantizer is inserted in the feedback path. Assume now $y[n]$ is rounded to 2 decimal places by the quantizer and the output is $y_Q[n]$. The input and the quantized output for $n = 0, 1, \dots, 15$ are tabulated in Table 7.9. We can see that $y_Q[n]$ decays into an oscillation of magnitude 0.01 after the input signal is turned off. The quantizers in the practical implementation of digital filters are nonlinear elements. Such nonlinearities in the feedback path may cause a recursive filter which is stable under infinite precision, to exhibit an unstable behavior under finite precision arithmetic for certain input signals. FIR filters are free from limit cycle oscillations, since they have no feedback loops.

Figure 7.10 A first order IIR filter with a quantizer in the feedback loop.

Table 7.9 The impulse response of the first order IIR filter shown in Figure 7.10.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$x[n]$	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$y_Q[n]$	1	-0.70	0.49	-0.34	0.24	-0.17	0.12	-0.08	0.06	-0.04	0.03	-0.02	0.01	-0.01	0.01	-0.01

Limit cycle oscillation is expected in our IIR compensation filter. We shall use the IIR filters designed in Example 6.1 to illustrate the limit cycle effect. Suppose an impulse signal is applied at the filter input. For the IIR-1 filter in Table 6.3, it is found that the output decays to a sustained oscillation of period of 20 and the sequence within one period is

$$\Delta \cdot (0, -3, 5, 6, -4, 0, 1, -1, -1, 1, 0, -1, 2, 3, -3, 0, 2, -4, -6, 5)$$

where Δ is output quantization step size. The magnitude of limit cycle oscillation is in the scale of the quantization step size. For the IIR-2 filter in Table 6.4, with the same input signal, the output decays to a repeated sequence of

$$\Delta \cdot (0, -3, 4, 6, -5, 0, 3, -5, -7, 5)$$

For the IIR-2 filter in Table 6.5, the output decays to zero.

Limit cycle oscillation may cause serious problem if it is not removed and if the subsequent stage has very high gain at the oscillation frequency. One method for removing limit cycle is by implementing the multiplier output with longer wordlength which makes the limit cycle amplitude acceptably low and then discarding those less significant bits associated with limit cycle. Since for an IIR filter, the internal processing wordlength is longer than the output wordlength to account for the higher noise, the limit cycle may be automatically removed during output wordlength truncation. Another method is to use magnitude truncation for the quantization arithmetic. Magnitude truncation selects the nearest quantized value that has a magnitude smaller than or equal to the magnitude of the unquantized value, i.e., truncating toward zero. The error caused by magnitude truncation, $e_Q[n]$, is in the range of $-\Delta < e_Q[n] < \Delta$, where Δ is the quantization step size. Magnitude quantization introduces more roundoff noise than rounding and truncation (truncating towards negative infinity) arithmetic. The noise power for magnitude truncation is $\Delta^2/3$ compared to $\Delta^2/12$ for rounding and truncation arithmetic.

7.6 Chapter Summary

In this chapter, the finite wordlength effects in our M -PTV IIR compensation filter. Three types of finite wordlength effects have been studied, namely, the roundoff noise, coefficient quantization and the limit cycle oscillation.

A white noise with zero mean appearing at the input of the M -PTV IIR filter becomes a zero mean noise with periodic time-varying variance at the filter output. For the roundoff noise contributed by input quantization, the noise amplification by the IIR filter is about the same as that by the FIR filter and is close to unity when the mismatch errors are relatively small. The roundoff noise contributed by multiplier output truncation is not amplified by the FIR compensation filter whereas it is amplified by the IIR compensation filter by a large factor. Longer wordlength is required for multiplier output truncation so that the total roundoff noise is not amplified by a big factor.

Coefficient quantization affects the filter response and hence the compensation performance. The IIR filters designed with smaller values of p_{max} are more robust under finite wordlength condition in the sense that their performance degrades by a smaller amount as the wordlength is decreased. IIR filters which achieve the same compensation performance as that of the FIR filters' under infinite wordlength condition may need longer coefficient wordlength to have the same performance as that of the FIR filter's under finite wordlength condition.

Our IIR compensation filter may suffer from limit cycle oscillations and the magnitude of the limit cycle is in the scale of the output quantization step size. Limit cycle may be removed by implementing the multiplier output with longer wordlength and then truncating those low order bits associated with limit cycle. Another method to remove the limit cycle is to use magnitude truncation.

As discussed in Chapter 6, when the FIR and the IIR filters are designed to achieve the same compensation performance, the IIR filters require 30% to 35% less number of multipliers than the FIR filters. On the other hand, the IIR filters need longer wordlength for multiplier output truncation and coefficient quantization. The implementation complexity of the IIR filters may be increased by the longer wordlength required. However, this is considered to be a very small amount of increase. In digital filter implementation, multiplications are the operations which consume most time, area and power. Moreover, if online filter design is required, the computational effort is directly related to the number of multipliers. Hence, compared to the resources saved by the 30% less multipliers, the resources increased by the longer wordlength is a very trivial part.

CHAPTER 8 : CONCLUSION AND FUTURE WORK

8.1 Conclusion

TIADCs achieve high sampling rate by employing a parallel array of slow sampling ADCs. The drawback associated with the TIADC structure is that any mismatch among the channels will introduce unwanted distortion images in the output spectrum and degrade the TIADC performance. Post-processing the TIADC output codes in the digital domain has attracted much research attention. Traditionally, FIR filter were employed in various post-processing structures to compensate for the mismatch errors. In this thesis, a novel structure which consists of recursive filters for TIADC mismatch compensation was demonstrated.

The basic idea of this new method is to estimate the ideal value of $x_m[k]$ using a weighted combination of the raw outputs from all the sub-ADCs sampled consecutively in the neighbourhood of $x_m[k]$ as well as the past estimated outputs which are closest to $x_m[k]$ where $x_m[k]$ is the k -th sample of the m -th sub-ADC. For an M -channel TIADC with overall sampling rate of f_s , our compensation structure consists of M IIR filters operating at interleaved time instants at the rate of f_s . The structure may be interpreted as multichannel IIR filters with M input ports or an M -PTV IIR filter. In Chapter 3, the basic design principle of the IIR mismatch compensation filter was demonstrated and a WLS optimization method for obtaining the filter coefficients was presented.

The stability of the proposed compensation system is of great concern. By using the state space representation of the IIR filter and applying vector processing, the M -PTV IIR compensation filter can be transformed into an equivalent M -input M -output LTI MIMO system operating at the channel

sampling rate f_s/M [133]. The MIMO system and hence the M -PTV IIR compensation filter is stable if all the eigenvalues of the state transition matrix of the MIMO system have magnitudes less than one. Utilizing the MIMO system, it can be further shown that the M -PTV IIR filter can be represented by a filter bank structure [134]. The equivalent filter bank consists of M parallel LTI IIR filters sharing a common denominator. The poles of the M -PTV IIR filter are the poles of the equivalent filter bank. Chapter 4 provided the detailed derivation procedures for obtaining the equivalent MIMO system from the M -PTV IIR filter and demonstrated that the pole locations of the M -PTV IIR filter depend solely on the denominator coefficients. The detailed derivation of the equivalent filter bank representation was also provided in Chapter 4. An alternative approach without using the state space representation for deriving the transfer function matrix of a different MIMO system which is operating at the rate of f_s was presented.

It is difficult to incorporate the eigenvalue based stability criterion in the WLS optimization process. Hence in Chapter 5 we introduced two iterative methods for gradually achieving a stable IIR filter design. One is by increasing the integer group delay η of the IIR compensation filter. Theoretical proof of this method by using *Rouché's Theorem* was presented. The other method is by adding positively weighted versions of the squared denominator coefficients into the WLS objective function. As the weights approach infinity, the magnitudes of the denominator coefficients approach zero and the filter becomes an FIR filter. Thus stability can always be achieved when the weights are sufficiently large. In Chapter 5, we have also investigated the minimum value of η required to achieve a stable design under various combinations of the design parameters, including M , N_a , N_b and the scale of the mismatch error ε_{max} . Two scenarios were considered, one is when only timing mismatch is present and the other is when only bandwidth mismatch is present. The results serve as useful guidelines for selecting a proper value of η to achieve a stable IIR compensation filter design.

Chapter 6 presented some design examples. It was shown that IIR compensation filters with guaranteed stability can be designed and the IIR filters have much lower implementation complexity than the FIR filters in [111] if they are designed to achieve the same level of compensation performance. Specifically, the multiplier saving is about 30% regardless of the scale of the mismatch

error. The number of delay elements and the number of arithmetic operations per sample are also about 30% less than those in the FIR case. The performance of our IIR filter in the presence of identification errors has been investigated and was found to be about the same as that of the FIR compensation filter. The compensation filters have to be redesigned to cope with the drift in the mismatch parameters. It was found through simulation examples that the stability and performance of the redesigned IIR compensation filters using the same design parameters are still guaranteed if the mismatch drift is not catastrophic and the poles of the initially designed filters are not too close to the unit circle. From our experience, if the design parameters are selected such that the maximum pole radius of the initially designed filter are less than 0.95, the stability of the redesigned filter is guaranteed for $r_{dr} \leq 0.7$. Two methods for pipeline implementation of the IIR filters to increase the operating speed were introduced. Both of these methods have increased implementation complexity compared to the original filter.

Finite wordlength effects of the IIR compensation filter have been discussed in Chapter 7. By using the equivalent MIMO representation, the noise propagation in the IIR filters was investigated. For the roundoff noise contributed by input quantization, the noise amplification by the IIR filter is about the same as that by the FIR filter and is close to unity for reasonably small mismatches. The roundoff noise contributed by the multiplier output truncation is not amplified by the FIR filter whereas it can be amplified by the IIR filter by a large factor. Coefficient quantization affects the filter response and hence the compensation performance of the IIR filters. It was found that the performances of the IIR filters designed with smaller pole radius degrade by a smaller amount when the coefficient wordlength decreases. Limit cycle effects may appear in the IIR compensation filter structure. IIR compensation filters need longer multiplier truncation wordlength to reduce the truncation noise and remove the limit cycle. However this is only a small amount of increase in the hardware resources compared to that saved by the 30% less multipliers.

8.2 Future Work

More comprehensive study of the finite wordlength effects in the IIR compensation filter can be carried out. Due to the time constrain, some aspects of the finite wordlength effects were not covered in this work. The internal nodes of the digital filters may suffer from overflow. For two's complement, overflow does not matter as long as the final sum is within the range represented by the finite wordlength. Overflow in IIR filters can cause another type of limit cycle oscillation. The amplitude of the oscillations can be as large as the dynamic range of the register experiencing the overflow and thus overflow oscillations are much more serious in nature. It is therefore important to scale the internal signal properly to avoid overflow.

So far, all the results obtained are based on simulations. In the future, the performance of the IIR compensation filter shall be tested using real life TIADC data. For some real-time applications where system stoppage is not tolerable, online redesign of the IIR compensation filter to cope with the mismatch drift is necessary. The actual hardware implementation of the IIR compensation filter and the filter redesign algorithm is another area to be explored.

The implementation complexity of the IIR compensation filter may be further reduced by using multiplier-less design where each filter coefficient is expressed as a sum of finite number of signed power of twos. From the implementation point of view, multiplications are then achieved by shifting the binary numbers and hence the design is multiplier-less.

Author's Publications

Part of the research work reported in this thesis are published or submitted for reviewing and are listed as follows.

1. S. Xu, Y. C. Lim, and J. W. Lee, "Recursive filters for time-interleaved ADC mismatch compensation", accepted by *IEEE Trans. Circuits Syst. I, Reg. Papers* for publication.
2. Y. C. Lim, S. Xu, and J. W. Lee, "Recursive filters for compensating the mismatch in time-interleaved ADCs", in *Proc. IEEE Int. Conf. Digital Signal Processing*, Singapore, 2015, pp. 324-328.
3. S. Xu, Y. C. Lim, J. H. Wong, and Q. H. Lam, "Massively parallel time-stretched analog-to-digital converter", in *Proc. 2013 Constantinides Int. Workshop Signal Processing*, London, 2013, pp. 1-4.
4. S. Xu and Y. C. Lim, "A review on time-interleaved analog-to-digital converters and mismatch compensation techniques", in "*Trends in Digital Signal Processing: A Festschrift in Honour of A. G. Constantinides*", Singapore: Pan Stanford Publishing Pte. Ltd., 2015, Chapter 2, pp. 13-63.

Bibliography

- [1] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp 539-550, Apr. 1999.
- [2] W. C. Black, Jr. and D. A. Hodges, "Time interleaved converter arrays," *IEEE J. Solid-State Circuits*, vol. SC-15, no. 6, pp. 1022-1029, Dec. 1980.
- [3] *IEEE standard for terminology and test methods for analog-to-digital converters*, IEEE Standard 1241-2000, Jun. 2001.
- [4] N. Kurosawa, H. Kobayashi, K. Maruyama, H. Sugawara, and K. Kobayashi, "Explicit analysis of channel mismatch effects in time-interleaved ADC systems," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, pp. 261-271, Mar. 2001.
- [5] C. Vogel, "The impact of combined channel mismatch effects in time-interleaved ADCs," *IEEE Trans. Instrum. Meas.*, vol. 54, pp. 415-427, Feb. 2005.
- [6] B. Yu and W. C. Black, Jr., "Error analysis for time-interleaved analog channels," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Sydney, 2001, pp. 468-471.
- [7] M. Gustavsson, J. J. Wikner, and N. Tan, *CMOS Data Converters for Communications*. Norwell, MA: Kluwer, 2000.
- [8] Y.-C. Jenq, "Digital spectra of nonuniformly sampled signals: fundamentals and high-speed waveform digitizers," *IEEE Trans. Instrum. Meas.*, vol. 37, pp. 245-251, Jun. 1988.
- [9] Y.-C. Jenq, "Digital spectra of nonuniformly sampled signals: a robust sampling time offset estimation algorithm for ultra high-speed waveform digitizers using interleaving," *IEEE Trans. Instrum. Meas.*, vol. 39, pp. 71-75, Feb. 1990.
- [10] A. Petraglia and S. K. Mitra, "Analysis of mismatch effects among A/D converters in a time-interleaved waveform digitizer," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 831-835, Oct. 1991.
- [11] S.-W. Sin, U.-F. Chio, S.-P. U, and R. P. Martins, "Statistical spectra and distortion analysis of

- time-interleaved sampling bandwidth mismatch,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, pp. 648-652, Jul. 2008.
- [12] C. S. G. Conroy, D. W. Cline, and P. R. Gray, “An 8-b 85-MS/s parallel pipeline A/D converter in 1- μ m CMOS,” *IEEE J. Solid-State Circuits*, vol. 28, no.4, pp. 447-454, Apr. 1993.
- [13] K. Y. Kim, N. Kusayanagi, and A. A. Abidi, “A 10-b, 100-MS/s CMOS A/D converter,” *IEEE J. Solid-State Circuits*, vol. 32, no.3, pp. 302-311, Mar. 1997.
- [14] L. Sumanen, M. Waltari, and K. A. I. Halonen, “A 10-bit 200-MS/s CMOS parallel pipeline A/D converter,” *IEEE J. Solid-State Circuits*, vol. 36, no.7, pp. 1048-1055, Jul. 2001.
- [15] K. Nakamura, M. Hotta, L. R. Carley, and D. J. Allstot, “An 85 mW, 10 b, 40 Msample/s CMOS parallel-pipelined ADC,” *IEEE J. Solid-State Circuits*, vol. 30, no. 3, pp. 173-183, Mar. 1995.
- [16] V. Ferragina, A. Fornasari, U. Gatti, P. Malcovati, and F. Maloberti, “Gain and offset mismatch calibration in time-interleaved multipath A/D sigma-delta modulators,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, pp. 2365-2373, Dec. 2004.
- [17] K. C. Dyer, D. Fu, S. H. Lewis, and P. J. Hurst, “An analog background calibration technique for time-interleaved analog-to-digital converters,” *IEEE J. Solid-State Circuits*, vol. 33, no.12, pp. 1912-1919, Dec. 1998.
- [18] K. C. Dyer, D. Fu, P. Hurst, and S. Lewis, “A comparison of monolithic background calibration in two time-interleaved analog-to-digital converters,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Monterey, 1998, pp. 13-16.
- [19] D. Fu, K. C. Dyer, S. H. Lewis, and P. J. Hurst, “A digital background calibration technique for time-interleaved analog-to-digital converters,” *IEEE J. Solid-State Circuits*, vol. 33, no.12, pp. 1904-1911, Dec. 1998.
- [20] J.-E. Eklund and F. Gustafsson, “Digital offset compensation of time-interleaved ADC using random chopper sampling,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Geneva, 2000, pp. 447-450.
- [21] A. Cabrini, F. Maloberti, R. Rovatti, and G. Setti, “On-line calibration of offset and gain mismatch in time-interleaved ADC using a sampled-data chaotic bit-stream,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Island of Kos, 2006, pp. 3398-3401.

- [22] S. M. Jamal, D. Fu, N. C.-J. Chang, P. J. Hurst, and S. H. Lewis, "A10-b 120-Msample/s time-interleaved analog-to-digital converter with digital background calibration," *IEEE J. Solid-State Circuits*, vol. 37, no. 12, pp. 1618-1627, Dec. 2002.
- [23] W. Liu and Y. Chiu, "Time-interleaved analog-to-digital conversion with online adaptive equalization," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 7, pp. 1384-1395, Jul. 2012.
- [24] A. Haftbaradaran and K. W. Martin, "Mismatch compensation techniques using random data for time-interleaved A/D converters," *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Island of Kos, 2006, pp. 3402-3405.
- [25] Y. Oh and B. Murmann, "System embedded ADC calibration for OFDM receivers," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, pp. 1693-1703, Aug. 2006.
- [26] J. M. D. Pereira, P. M. B. S. Girao, and A. M. C. Serra, "An FFT-based method to evaluate and compensate gain and offset errors of interleaved ADC systems," *IEEE Trans. Instrum. Meas.*, vol. 53, pp. 423-430, Apr. 2004.
- [27] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Blind adaptive equalization of mismatch errors in a time-interleaved A/D converter system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, pp. 151-158, Jan. 2004.
- [28] C. H. Law, P. J. Hurst and S. H. Lewis, "A four-channel time-interleaved ADC with digital calibration of interchannel timing and memory errors," *IEEE J. Solid-State Circuits*, vol. 45, no. 10, pp. 2091-2103, Oct. 2010.
- [29] J. Matsuno, T. Yamaji, M. Furuta, and T. Itakura, "All-digital background calibration technique for time-interleaved ADC using pseudo aliasing signal," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, pp. 1113-1121, May 2013.
- [30] K. El-Sankary and M. Sawam, "A background calibration technique for multibit/stage pipelined and time-interleaved ADCs," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 6, pp. 448-452, Jun. 2006.
- [31] H. Jin, E. Lee, and M. Hassoun, "Time-interleaved A/D converter with channel randomization," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Hong Kong, 1997, pp. 425-428.
- [32] G. Bemardinis, P. Malcovati, F. Maloberti, and E. Soenen, "Dynamic Stage Matching for

- Parallel Pipeline A/D Converters,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Phoenix-Scottsdale, AZ, 2002, pp. 905- 908.
- [33] K. Poulton, J. J. Corcoran, and T. Hornak, “A 1-GHz 6-bit ADC system,” *IEEE J. Solid-State Circuits*, vol. SC-22, no.6, pp. 962-970, Dec. 1987.
- [34] S. Gupta, M. Choi, M. Interfield, and J. Wang, “A 1 GS/s 11b time-interleaved ADC in 0.13 μm CMOS,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, San Francisco, CA, 2006, pp. 2360-2369.
- [35] S. K. Gupta, M. A. Inerfield, and Jingbo Wang, “A 1-GS/s 11-bit ADC with 55-dB SNDR, 250-mW power realized by a high bandwidth scalable time-interleaved architecture,” *IEEE J. Solid-State Circuits*, vol. 41, no. 12, Dec. 2006.
- [36] K. Poulton, K. L. Knudsen, J. Kerley, J. Kang, J. Tani, E. Cornish, and M. VanGrouw, “An 8-GSa/s 8-bit ADC system,” in *Symp. VLSI Circuits Dig. Tech. Papers*, Kyoto, 1997, pp. 23-24.
- [37] K. Poulton, R. Neff, A. Muto, W. Liu, A. Burstein, and M. Heshami, “A 4 GSample/s 8b ADC in 0.35 μm CMOS,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, San Francisco, CA, 2006, pp. 166-457.
- [38] K. Poulton, R. Neff, B. Setterberg, B. Wuppermann, T. Kopley, R. Jewett, J. Pernillo, C. Tan, and A. Montijo, “A 20 GS/s 8b ADC with a 1 MB memory in 0.18 μm CMOS,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, San Francisco, CA, 2003, pp. 318-496.
- [39] A. Haftbaradaran and K. W. Martin, “A background compensation technique for sample-time errors in time-interleaved A/D converters,” in *Proc. of the IEEE Int. Midwest Symp. Circuits and Systems (MWSCAS)*, Covington, 2005, pp. 1011-1014.
- [40] A. Haftbaradaran and K. W. Martin, “A sample-time error compensation technique for time-interleaved ADC systems,” in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, San Jose, CA, 2007, pp. 341-344.
- [41] A. Haftbaradaran and K. W. Martin, “A background sample-time error calibration technique using random data for wide-band high-resolution time-interleaved ADCs,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, pp. 234-238, Mar. 2008.
- [42] C.-Y. Wang and J.-T. Wu, “A background timing-skew calibration technique for time-interleaved analog-to-digital converters,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, pp.

- 299-303, Apr. 2006.
- [43] C.-Y. Wang and J.-T. Wu, "A multiphase timing-skew calibration technique using zero-crossing detection," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 1102-1114, Jun. 2009.
 - [44] C.-C. Huang, C.-Y. Wang and J.-T. Wu, "A CMOS 6-bit 16-GS/s time-interleaved ADC using digital background calibration techniques," *IEEE J. Solid-State Circuits*, vol. 46, no. 4, pp. 848-858, Apr. 2011.
 - [45] M. El-Chammas and B. Murmann, "A 12-GS/s, 81-mW 5-bit time-interleaved flash ADC with background timing skew calibration," *IEEE J. Solid-State Circuits*, vol. 46, no. 4, pp. 838-847, Apr. 2011.
 - [46] T.-Y. Tang, T.-H. Tsai, and K. Chen, "Timing mismatch background calibration for time-interleaved ADCs," in *Proc. IEEE Region 10 Conf. TENCN*, Cebu, 2012, pp. 1-4.
 - [47] B. Razavi, "Design considerations for interleaved ADCs," *IEEE J. Solid-State Circuits*, vol. 48, no. 8, pp. 1806-1817, Aug. 2013.
 - [48] J. Li *et al.*, "A digital timing mismatch calibration technique in time-interleaved ADCs," *IEEE Trans. Circuits Sys. II: Exp. Briefs*, vol. 61, pp. 486-490, Jul. 2014.
 - [49] C. Vogel, D. Draxelmayr and F. Kuttner, "Compensation of timing mismatches in time-interleaved analog-to-digital converters through transfer characteristic tuning," in *Proc. IEEE Int. Midwest Symp. Circuits and Systems (MWSCAS)*, Hiroshima, 2004, pp. 341-344.
 - [50] K. Yao and J. Thomas, "On some stability and interpolatory properties of nonuniform sampling expansions," *IEEE Trans. Circuits Theory*, vol. CT-14, pp. 404-408, Dec. 1967.
 - [51] P. P. Vaidyanathan, "Maximally decimated filter banks" in *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993, ch. 5, pp. 188-285.
 - [52] Y. C. Eldar and A. V. Oppenheim, "Filterbank reconstruction of bandlimited signals from nonuniform and generalized samples," *IEEE Trans. Signal Process.*, vol. 48, pp. 2864-2875, Oct. 2000.
 - [53] H. Johansson and P. Löwenborg, "Reconstruction of nonuniformly sampled bandlimited signals by means of digital fractional delay filters," *IEEE Trans. Signal Process.*, vol. 50, pp. 2757-2767, Nov. 2002.

- [54] H. Johansson and P. Löwenborg, "Reconstruction of nonuniformly sampled bandlimited signals using digital fractional delay filters," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Sydney, 2001, pp. 593-596.
- [55] H. Johansson and P. Löwenborg, "Reconstruction of a class of nonuniformly sampled and decimated bandlimited signals," in *Proc. IEEE Int. Symp. Circuits Syst., (ISCAS)*, Phoenix-Scottsdale, AZ, 2002, pp. 604-607.
- [56] R. S. Prendergast, B. C. Levy, and P. J. Hurst, "Reconstruction of band-limited periodic nonuniformly sampled signals through multirate filter banks," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, pp. 1612-1622, Aug. 2004.
- [57] W. Namgoong, "Finite-length synthesis filters for non-uniformly time-interleaved analog-to-digital converter," in *Proc. IEEE Int. Symp. Circuits Syst., (ISCAS)*, Phoenix-Scottsdale, AZ, 2002, pp. 815-818.
- [58] D. Marelli, K. Mahata, and M. Fu, "Linear LMS compensation for timing mismatch in time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 2476-2486, Nov. 2009.
- [59] H. Johansson and P. Löwenborg, "Reconstruction of nonuniformly sampled bandlimited signals by means of time-varying discrete-time FIR filters," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1-18, 2006, Article ID 64185.
- [60] H. Johansson, P. Löwenborg, and K. Vengattaramane, "Least-squares and minimax design of polynomial impulse response FIR filters for reconstruction of two-periodic nonuniformly sampled signals," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, pp. 877-888, Apr. 2007.
- [61] H. Johansson, P. Löwenborg, and K. Vengattaramane, "Reconstruction of M -periodic nonuniformly sampled signals using multivariate polynomial impulse response time-varying FIR filters," in *Proc. European Signal Process. Conf.*, Florence, 2006, pp. 1-4.
- [62] A. K. M. Pillai and H. Johansson, "Efficient signal reconstruction scheme for M -channel time-interleaved ADCs," *Analog Integrated Circuits and Signal Process.*, vol. 77, no. 2, pp. 113-122, Nov. 2013.
- [63] A. K. M. Pillai and H. Johansson, "Low-complexity two-rate based multivariate impulse response reconstructor for time-skew error correction in M -channel time-interleaved ADCs," in

- Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Beijing, 2013, pp. 2936-2939.
- [64] S. Tertinek and C. Vogel, "Reconstruction of two-periodic nonuniformly sampled band-limited signals using a discrete-time differentiator and a time-varying multiplier," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 54, pp. 616-620, Jul. 2007.
- [65] S. Tertinek and C. Vogel, "Reconstruction of nonuniformly sampled bandlimited signals using a differentiator-multiplier cascade," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, pp. 2273-2286, Sep. 2008.
- [66] H. Le Duc *et al.*, "A fully digital background calibration of timing skew in undersampling TI-ADC," in *Proc. IEEE 12th Int. New Circuits System Conf. (NEWCAS)*, Trois-Rivieres, 2014, pp. 53-56.
- [67] H. Le Duc *et al.*, "Hardware implementation of all digital calibration for undersampling TIADCs," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Lisbon, 2015, pp. 2181-2184.
- [68] H. L. Duc *et al.*, "All-digital calibration of timing skews for TIADCs using the polyphase decomposition," *IEEE Trans. Circuits Syst. II: Express briefs*, vol. 63, pp. 99-104, Jan. 2016.
- [69] N. Le Dortz *et al.*, "A 1.62GS/s time-interleaved SAR ADC with digital background mismatch calibration achieving interleaving spurs below 70dBFS," in *IEEE Int. Solid-State Circuits Conf. Digest. Technical Paper (ISSCC)*, San Francisco, CA, 2014, pp. 386-388.
- [70] F. Centurelli, P. Monsurrò, and A. Trifiletti, "Efficient digital background calibration of time-interleaved pipeline analog-to-digital converters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, pp. 1373-1383, Jul. 2012.
- [71] V. Divi and G. Wornell, "Bandlimited signal reconstruction from noisy periodic nonuniform samples in time-interleaved ADCs," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, Nevada, 2008, pp. 3721-3724.
- [72] K. M. Tsui and S. C. Chan, "A versatile iterative framework for the reconstruction of bandlimited signals from their nonuniform sample," *J. Signal. Process. Syst.*, vol. 62, no. 3, pp. 459-468, Mar. 2011.
- [73] H. Jin and K. F. Lee, "A digital-background calibration technique for minimizing timing-error effects in time-interleaved ADC's," *IEEE Trans. Circuits Syst. II*, vol. 47, no.7, pp. 603-613,

Jul. 2000.

- [74] Y. X. Zou, S. L. Zhang, Y. C. Lim, and X. Chen, "Timing mismatch compensation in time-interleaved ADCs based on multichannel Lagrange polynomial interpolation," *IEEE Trans. Instrum. Meas.*, vol. 60, pp. 1123-1131, Apr. 2011.
- [75] C. A. Schmidt *et al.*, "Efficient estimation and correction of mismatch errors in time-interleaved ADCs," *IEEE Trans. Instrum. Meas.*, vol. 65, pp. 243-254, Feb. 2016.
- [76] Y.-C. Jenq, "Perfect reconstruction of digital spectrum from nonuniformly sampled signals," *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 649-652, Jun. 1997.
- [77] Y.-C. Jenq, "Perfect reconstruction of digital spectrum from nonuniformly sampled signals," in *Proc. IEEE. Instrumentation and Measurement Technology Conf. (IMTC)*, Ottawa, 1997, pp. 624-627.
- [78] C. W. Farrow, "A continuously variable digital delay element," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Espoo, 1988, pp. 2641-2645.
- [79] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Analysis of mismatch noise in randomly interleaved ADC system," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Hong Kong, 2003, pp. 277-280.
- [80] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Analysis of mismatch effects in a randomly interleaved A/D converter system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, pp. 465-476, Mar. 2005.
- [81] M. Tamba, A. Shimizu, H. Munakata, and T. Komuro, "A method to improve SFDR with random interleaved sampling method," in *Proc. Int. Test Conf. (ITC)*, Baltimore, MD, 2001, pp. 512-520.
- [82] C. Vogel, D. Draxelmayr, and G. Kubin, "Spectral shaping of timing mismatches in time-interleaved analog-to-digital converters," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Kobe, 2005, pp. 1394-1397.
- [83] C. Vogel, V. Pammer, and G. Kubin, "A novel channel randomization method for time-interleaved ADCs," in *Proc. IEEE. Instrumentation and Measurement Technology Conf. (IMTC)*, Ottawa, 2005, pp. 150-155.
- [84] Y.-C. Jenq, "Improving timing offset estimation by alias sampling," *IEEE Trans. Instrum.*

- Meas.*, vol. 57, pp. 1376-1378, Jul. 2008.
- [85] E. Iroaga, B. Murmann, and L. Nathawad, "A background correction technique for timing errors in time-interleaved analog-to-digital converters," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Kobe, 2005, pp. 5557-5560.
 - [86] V. Divi and G. Wornell, "Signal recovery in time-interleaved analog-to-digital converters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Montreal, 2004, pp. 593-596.
 - [87] V. Divi and G. Wornell, "Scalable blind calibration of timing skew in high-resolution time-interleaved ADCs," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Island of Kos, 2006, pp. 3390-3393.
 - [88] V. Divi and G. W. Wornell, "Blind calibration of timing skew in time-interleaved analog-to-digital converters," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 3, pp. 509-522, Jun. 2009.
 - [89] S. Huang and B. C. Levy, "Adaptive blind calibration of timing offset and gain mismatch for two-channel time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, pp. 1278-1288, Jun. 2006.
 - [90] S. Huang and B. C. Levy, "Blind calibration of timing offsets for four-channel time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, pp. 863-876, Apr. 2007.
 - [91] C. Vogel, "A frequency domain method for blind identification of timing mismatches in time-interleaved ADCs," in *Proc. Norchip Conf.*, Linkoping, 2006, pp. 45-48.
 - [92] S. Saleem and C. Vogel, "LMS-based identification and compensation of timing mismatches in a two-channel time-interleaved analog-to-digital converter," in *Proc. Norchip Conf.*, Aalborg, 2007, pp. 1-4.
 - [93] C. Vogel, S. Saleem, and S. Mendel, "Adaptive blind compensation of gain and timing mismatches in M -channel time-interleaved ADCs," in *Proc. IEEE Int. Conf. Electronics, Circuits and Systems (ICECS)*, St. Julien's, 2008, pp. 49-52.
 - [94] S. Saleem and C. Vogel, "On blind identification of gain and timing mismatches in time-interleaved analog-to-digital converters," in *Int. Conf. Telecommunications and Signal Processing (TSP)*, Baden, 2010, pp. 151-155.
 - [95] Y. X. Zou and X. J. Xu, "Blind timing skew estimation using source spectrum sparsity in time-

- interleaved ADCs”, *IEEE Trans. Instrum. Meas.*, vol. 61, pp. 2401-2412, Sep. 2012.
- [96] J. Elbornsson and J.-E. Eklund, “Blind estimation of timing errors in interleaved AD converters,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Salt Lake City, 2001, pp. 3913-3916.
- [97] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, “Amplitude and gain error influence on time error estimation algorithm for time interleaved A/D converter system,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Orlando, FL, 2002, pp. 1281-1284.
- [98] J. Elbornsson, K. Folkesson, and J.-E. Eklund, “Measurement verification of estimation method for time errors in a time interleaved A/D converter system,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Phoenix-Scottsdale, AZ, 2002, pp. 129-132.
- [99] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, “Blind equalization of time errors in a time-interleaved ADC system,” *IEEE Trans. Signal Process.*, vol. 53, pp. 1413-1424, Apr. 2005.
- [100] C. Luo, L. Zhu, and J. McClellan, “Coordinated blind calibration for time interleaved ADCs,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, 2013, pp. 3890-3894.
- [101] M. Seo, M. J. W. Rodwell, and U. Madhow, “Blind correction of gain and timing mismatches for a two-channel time-interleaved analog-to-digital converter,” in *Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, 2005, pp. 1121-1125.
- [102] M. Seo, M. J. W. Rodwell, and U. Madhow, “Blind correction of gain and timing mismatches for a two-channel time-interleaved analog-to-digital converter: experimental verification,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, Island of Kos, 2006, pp. 3394-3397.
- [103] S. M. Jamal, D. Fu, M. P. Singh, P. J. Hurst, and S. H. Lewis, “Calibration of sample-time error in a two-channel time-interleaved analog-to-digital converter,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, pp. 130-139, Jan. 2004.
- [104] T.-H. Tsai, P. J. Hurst, and S. H. Lewis, “Bandwidth mismatch and its correction in time-interleaved analog-to-digital converters,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, pp. 1133-1137, Oct. 2006.
- [105] P. Satarzadeh, B. C. Levy, and P. J. Hurst, “Bandwidth mismatch correction for a two-channel time-interleaved A/D converter,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, New

- Orleans, LA, 2007, pp. 1705-1708.
- [106] P. Satarzadeh, B. C. Levy and P. J. Hurst, "Adaptive semiblind calibration of bandwidth mismatch for two-channel time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 2075-2088, Sep. 2009.
 - [107] T.-H. Tsai, P. J. Hurst, and S. H. Lewis, "Correction of mismatches in a time-interleaved analog-to-digital converter in an adaptively equalized digital communication receiver," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 307-319, Feb. 2009.
 - [108] K. Asami, "Technique to improve the performance of time-interleaved A-D converters," in *Proc. Int. Test Conf. (ITC)*, Austin, TX, 2005, pp. 1-7.
 - [109] C. Cho *et al.*, "Calibration of channel mismatch in time-interleaved real-time digital oscilloscopes," in *Proc. 25th IEEE Microwave Measurement Conf.*, Phoenix, AZ, 2015, pp. 1-5.
 - [110] M. Seo, M. J. W. Rodwell, and U. Madhow, "Comprehensive digital correction of mismatch errors for a 400-Msamples/s 80-dB SFDR time-interleaved analog-to-digital converter," *IEEE Trans. Microw. Theory Tech.*, vol. 53, pp. 1072-1082, Mar. 2005.
 - [111] Y. C. Lim, Y.-X. Zou, J. W. Lee, and S.-C. Chan, "Time-interleaved analog-to-digital-converter compensation using multichannel filters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 2234-2247, Oct. 2009.
 - [112] H. Johansson and P. Löwenborg, "A least-squares filter design technique for the compensation of frequency-response mismatch errors in time-interleaved A/D converters," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, pp. 1154-1158, Nov. 2008.
 - [113] S. Mendel and C. Vogel, "A compensation method for magnitude response mismatches in two-channel time-interleaved analog-to-digital converters," in *Proc. 13th IEEE Int. Conf. Electronics, Circuits and Systems*, Nice, 2006, pp. 712-715.
 - [114] S. Mendel and C. Vogel, "On the compensation of magnitude response mismatches in M -channel time-interleaved ADCs," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, New Orleans, LA, 2007, pp. 3375-3378.
 - [115] C. Vogel and S. Mendel, "A flexible and scalable structure to compensate frequency response mismatches in time-interleaved ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 2463-2475, Nov. 2009.

- [116] H. Johansson, "A polynomial-based time-varying filter structure for the compensation of frequency-response mismatch errors in time-interleaved ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 3, pp. 384-396, Jun. 2009.
- [117] K. M. Tsui and S. C. Chan, "New iterative framework for frequency response mismatch correction in time-interleaved ADCs: design and performance analysis," *IEEE Trans. Instrum. Meas.*, vol. 60, pp. 3792-3805, Nov. 2011.
- [118] K. M. Tsui and S. C. Chan, "Iterative correction of frequency response mismatches in time-interleaved ADCs: a novel framework and case study in OFDM systems," in *Proc. Int. Conf. Green Circuits and Systems (ICGCS)*, Shanghai, pp. 253-258.
- [119] M. Seo, M. Rodwell, and U. Madhow, "Generalized blind mismatch correction for two-channel time-interleaved A-to-D converters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, HI, 2007, pp. 1505-1508.
- [120] M. Seo and M. Rodwell, "Generalized blind mismatch correction for a two-channel time-interleaved ADC: analytic approach," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, New Orleans, LA, 2007, pp. 109-112.
- [121] S. Saleem and C. Vogel, "Adaptive compensation of frequency response mismatches in high-resolution time-interleaved ADCs using a low-resolution ADC and a time-varying filter," in *Proc. 2010 IEEE Int. Symp. On Circuits and Systems (ISCAS)*, Paris, 2010, pp 561-564.
- [122] S. Saleem and C. Vogel, "Adaptive blind background calibration of polynomial-represented frequency response mismatches in a two-channel time-interleaved ADC," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, pp. 1300-1310, Jun. 2011.
- [123] G. K. D. Teyou, H. Petit, and P. Loumeau, "Adaptive and digital blind calibration of transfer function mismatch in time-interleaved ADCs," in *Proc. IEEE 13th Int. New Circuits Systems Conf. (NEWCAS)*, Grenoble, 2015, pp. 1-4.
- [124] S. J. Liu *et al.*, "Adaptive calibration of channel mismatches in time-interleaved ADCs based on equivalent signal recombination," *IEEE Trans. Instrum. Meas.*, vol. 63, pp. 277-286, Feb. 2014.
- [125] A. Bonnetat *et al.*, "Correlation-based frequency-response mismatch compensation of quad-TIADC using real samples," *IEEE Trans. Circuits Syst. II, Exp Briefs*, vol. 62, pp. 746-750,

Aug. 2015.

- [126] A. Bonnetat *et al.*, “An adaptive all-digital blind compensation of dual-TIADC frequency-response mismatch based on complex signal correlations,” *IEEE Trans. Circuits Syst. II, Exp Briefs*, vol. 62, pp. 821-825, Sep. 2015.
- [127] Y. C. Lim, J.-H. Lee, C. K. Chen, and R.-H. Yang, “A weighted least squares algorithm for quasi-equiripple FIR and IIR digital filter design,” *IEEE Trans. Signal Process.*, vol. 40, pp. 551-558, Mar. 1992.
- [128] W.-S. Lu, S.-C. Pei, and C.-C. Tseng, “A weighted least squares method for the design of stable 1-D and 2-D IIR digital filters,” *IEEE Trans. Signal Process.*, vol. 46, pp. 1-10, Jan. 1998.
- [129] M. C. Lang, “Least-squares design of IIR filters with prescribed magnitude and phase responses and a pole radius constraint,” *IEEE Trans. Signal Process.*, vol. 48, pp. 3109-3121, Nov. 2000.
- [130] A. Jiang and H. K. Kwan, “IIR digital filter design with novel stability criterion based on argument principle,” in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, New Orleans, LA, 2007, pp. 2339-2342.
- [131] T. C. Hsia, *System Identification: Least-Squares Methods*, 1st ed. Lanham, Maryland: Lexington Books, 1997.
- [132] S. Burrus, “Block realization of digital filters,” *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 230-235, Oct. 1972.
- [133] R. A. Meyer and C. S. Burrus, “A unified analysis of multirate and periodically time-varying digital filters,” *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 162-168, Mar. 1975.
- [134] J. S. Prater and C. S. Burrus, “Analysis and design of periodically time-varying IIR filters, with applications to transmultiplexing,” *IEEE Trans. Signal Process.*, vol. 40, pp. 2715-2725, Nov. 1992.
- [135] J. H. Mathews and R. W. Howell, “Residue theory,” in *Complex Analysis for Mathematics and Engineering*, 4th ed., Sudbury: Jones and Bartlett Publishers, 2001, pp. 349.
- [136] J. H. Wong et al, “Photonic time-stretched analog-to-digital converter amenable to continuous-time operation based on polarization modulation with balanced detection scheme,” *J. Lightw.*

- Technol.*, vol. 29, no. 20, pp. 3099-3106, Oct. 2011.
- [137] H. H. Loomis, Jr. and B. Sinha, "High-speed recursive digital filter realization", *Circuits, Syst. and Signal Process.*, vol. 3, No. 3, pp. 267-294, Sep. 1984.
- [138] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, "Structures for discrete-time systems" in *Discrete-Time Signal Processing*, 2nd ed. New Jersey: Prentice-Hall, 1999, ch. 6, pp. 340-438.
- [139] S. K. Mitra, "Analysis of finite wordlength effects" in *Digital Signal Processing, A Computer-Based Approach*, 3rd ed. New York: McGraw Hill, 2006, ch.12, pp. 665-738.
- [140] D. Schlichtharle, "Effects of finite register lengths" in *Digital Filters: Basics and Design*, Heidelberg, Germany: Springer-Verlag, 2010, ch. 8, pp. 383-467.
- [141] F. J. Taylor, "Fixed-point effects" in *Digital Filters: Principles and Applications with MATLAB*, New York: Wiley-IEEE Press, 2012, ch. 16, pp. 215-230.
- [142] B. W. Bomar, "Finite wordlength effects" in *Digital Signal Processing Handbook*, Boca Raton: CRC Press LLC, 1999, ch. 3.

Appendix A: Detailed Derivation of the Equivalent MIMO System Using State Space Representation

The state space representation of the proposed M -periodic time-varying IIR filter is given by

$$\mathbf{s}[n+1] = \mathbf{A}_n \mathbf{s}[n] + \mathbf{b}_n x[n] \quad (\text{A1})$$

$$y[n] = \mathbf{c}_n^T \mathbf{s}[n] + d_n x[n] \quad (\text{A2})$$

The state space parameters $(\mathbf{A}_n, \mathbf{b}_n, \mathbf{c}_n, d_n)$ are M -PTV, i.e. $\mathbf{A}_n = \mathbf{A}_{n+M}$, $\mathbf{b}_n = \mathbf{b}_{n+M}$, $\mathbf{c}_n = \mathbf{c}_{n+M}$, and $d_n = d_{n+M}$.

Define a new state vector $\mathbf{t}[n]$ as the down-sampled version of $\mathbf{s}[n]$ such that $\mathbf{t}[n] = \mathbf{s}[nM]$. Using (A1) and the M -periodicity of $\mathbf{A}_n$ repeatedly, $\mathbf{t}[n+1]$ can be written as

$$\begin{aligned} \mathbf{t}[n+1] &= \mathbf{A}_{M-1} \mathbf{s}[nM+M-1] + \mathbf{b}_{M-1} x[nM+M-1] \\ &= \mathbf{A}_{M-1} \mathbf{A}_{M-2} \mathbf{s}[nM+M-2] + \mathbf{A}_{M-1} \mathbf{b}_{M-2} x[nM+M-2] + \mathbf{b}_{M-1} x[nM+M-1] \\ &= \mathbf{A}_{M-1} \mathbf{A}_{M-2} \mathbf{A}_{M-3} \mathbf{s}[nM+M-3] + \mathbf{A}_{M-1} \mathbf{A}_{M-2} \mathbf{b}_{M-3} x[nM+M-3] \\ &\quad + \mathbf{A}_{M-1} \mathbf{b}_{M-2} x[nM+M-2] + \mathbf{b}_{M-1} x[nM+M-1] \\ &= \dots \\ &= (\mathbf{A}_{M-1} \mathbf{A}_{M-2} \dots \mathbf{A}_0) \mathbf{s}[nM] + (\mathbf{A}_{M-1} \dots \mathbf{A}_1 \mathbf{b}_0) x[nM] + (\mathbf{A}_{M-1} \dots \mathbf{A}_2 \mathbf{b}_1) x[nM+1] \\ &\quad + \dots + (\mathbf{A}_{M-1} \mathbf{b}_{M-2}) x[nM+M-2] + (\mathbf{b}_{M-1}) x[nM+M-1] \end{aligned} \quad (\text{A3})$$

(A3) is rewritten in matrix-vector form as

$$\mathbf{t}[n+1] = \tilde{\mathbf{A}} \mathbf{t}[n] + \tilde{\mathbf{B}} x[n] \quad (\text{A4})$$

where

$$\mathbf{x}[n] = [x[nM], x[nM + 1], \dots, x[nM + M - 1]]^T, \quad (\text{A5})$$

$$\tilde{\mathbf{A}} = \mathbf{A}_{M-1} \mathbf{A}_{M-2} \dots \mathbf{A}_1 \mathbf{A}_0, \quad (\text{A6})$$

$$\tilde{\mathbf{B}} = [(\mathbf{A}_{M-1} \mathbf{A}_{M-2} \dots \mathbf{A}_1 \mathbf{b}_0), (\mathbf{A}_{M-1} \mathbf{A}_{M-2} \dots \mathbf{A}_2 \mathbf{b}_1) \dots, (\mathbf{A}_{M-1} \mathbf{b}_{M-2}), \mathbf{b}_{M-1}], \quad (\text{A7})$$

Define the output vector $\mathbf{y}[n]$ as

$$\mathbf{y}[n] = [y[nM], y[nM + 1], \dots, y[nM + M - 1]]^T. \quad (\text{A8})$$

Expand each element of $\mathbf{y}[n]$ in a similar manner and we have

$$\begin{aligned} y[nM] &= \mathbf{c}_0^T \mathbf{t}[n] + d_0 x[nM] \\ y[nM + 1] &= \mathbf{c}_1^T \mathbf{s}[nM + 1] + d_1 x[nM + 1] \\ &= \mathbf{c}_1^T \mathbf{A}_0 \mathbf{t}[n] + \mathbf{c}_1^T \mathbf{b}_0 x[nM] + d_1 x[nM + 1] \\ y[nM + 2] &= \mathbf{c}_2^T \mathbf{s}[nM + 2] + d_2 x[nM + 2] \\ &= \mathbf{c}_2^T \mathbf{A}_1 \mathbf{s}[nM + 1] + \mathbf{c}_2^T \mathbf{b}_1 x[nM + 1] + d_2 x[nM + 2] \\ &= \mathbf{c}_2^T \mathbf{A}_1 \mathbf{A}_0 \mathbf{t}[n] + \mathbf{c}_2^T \mathbf{A}_1 \mathbf{b}_0 x[nM] + \mathbf{c}_2^T \mathbf{b}_1 x[nM + 1] + d_2 x[nM + 2] \\ &\dots \\ y[nM + M - 1] &= \mathbf{c}_{M-1}^T \mathbf{s}[nM + M - 1] + d_{M-1} x[nM + M - 1] \\ &= \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{s}[nM + M - 2] + \mathbf{c}_{M-1}^T \mathbf{b}_{M-2} x[nM + M - 2] \\ &\quad + d_{M-1} x[nM + M - 1] \\ &= \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{A}_{M-3} \mathbf{s}[nM + M - 3] + \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{b}_{M-3} x[nM + M - 3] \\ &\quad + \mathbf{c}_{M-1}^T \mathbf{b}_{M-2} x[nM + M - 2] + d_{M-1} x[nM + M - 1] \\ &= \dots \\ &= (\mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{A}_{M-3} \dots \mathbf{A}_0) \mathbf{t}[n] + (\mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{A}_{M-3} \dots \mathbf{A}_1 \mathbf{b}_0) x[nM] \\ &\quad + (\mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{A}_{M-3} \dots \mathbf{A}_2 \mathbf{b}_1) x[nM + 1] \\ &\quad + \dots (\mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{b}_{M-3}) x[nM + M - 3] + (\mathbf{c}_{M-1}^T \mathbf{b}_{M-2}) x[nM + M - 2] \\ &\quad + d_{M-1} x[nM + M - 1] \end{aligned} \quad (\text{A9})$$

Combining the equations in (A9) in matrix form leads to

$$\mathbf{y}[n] = \tilde{\mathbf{C}}\mathbf{t}[n] + \tilde{\mathbf{D}}\mathbf{x}[n] \quad (\text{A10})$$

where

$$\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{c}_0^T \\ \mathbf{c}_1^T \mathbf{A}_0 \\ \mathbf{c}_2^T \mathbf{A}_1 \mathbf{A}_0 \\ \vdots \\ \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \mathbf{A}_{M-3} \cdots \mathbf{A}_0 \end{bmatrix} \quad (\text{A11})$$

and $\tilde{\mathbf{D}}$ is a lower triangular matrix and

$$\tilde{\mathbf{D}} = \begin{bmatrix} d_0 & 0 & 0 & \cdots & 0 & 0 \\ \mathbf{c}_1^T \mathbf{b}_0 & d_1 & 0 & \cdots & 0 & 0 \\ \mathbf{c}_2^T \mathbf{A}_1 \mathbf{b}_0 & \mathbf{c}_2^T \mathbf{b}_1 & d_2 & \cdots & 0 & 0 \\ \mathbf{c}_3^T \mathbf{A}_2 \mathbf{A}_1 \mathbf{b}_0 & \mathbf{c}_3^T \mathbf{A}_2 \mathbf{b}_1 & \mathbf{c}_3^T \mathbf{b}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \cdots \mathbf{A}_1 \mathbf{b}_0 & \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \cdots \mathbf{A}_2 \mathbf{b}_1 & \mathbf{c}_{M-1}^T \mathbf{A}_{M-2} \cdots \mathbf{A}_3 \mathbf{b}_2 & \cdots & \mathbf{c}_{M-1}^T \mathbf{b}_{M-2} & d_{M-1} \end{bmatrix} \quad (\text{A12})$$

The element of $\tilde{\mathbf{D}}$ at the i -th row and the j -th column $\tilde{d}_{ij}$ is given by, assuming the indices i and j start from zero

$$\tilde{d}_{ij} = \begin{cases} 0 & i < j \\ d_i & i = j \\ \mathbf{c}_i^T \mathbf{b}_j & i = j + 1 \\ \mathbf{c}_i^T \mathbf{A}_{i-1} \cdots \mathbf{A}_{j+1} \mathbf{b}_j & i \geq j + 2 \end{cases} \quad (\text{A13})$$

Appendix B: Eigenvalues of the Matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{A}}'$

Partition the state transition matrix of the M -PTV IIR filter, $\mathbf{A}_n$, given in (4.5) into the following four matrices

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{A}'_n & \mathbf{U}_n \\ \mathbf{V} & \mathbf{W} \end{bmatrix} \quad (\text{B1})$$

where

$$\mathbf{A}'_n = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -b_n[N_b] & -b_n[N_b - 1] & \dots & -b_n[1] \end{bmatrix} \quad (\text{B2})$$

$$\mathbf{U}_n = \begin{bmatrix} \mathbf{0}_{(N_b-1) \times (N_a-1)} \\ [a_n[N_a - 1], \dots, a_n[1]] \end{bmatrix} \quad (\text{B3})$$

$$\mathbf{V} = \mathbf{0}_{(N_a-1) \times N_b} \quad (\text{B4})$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{0}_{(N_a-2) \times 1} & \mathbf{I}_{(N_a-2)} \\ 0 & \mathbf{0}_{1 \times (N_a-2)} \end{bmatrix} \quad (\text{B5})$$

In (B2-B5), $\mathbf{0}_{m \times n}$ represents the null matrix/vector consisting of m rows and n columns and $\mathbf{I}_m$ represents the identity matrix with dimension of $(m \times m)$. The M -PTV IIR filter has an equivalent MIMO representation and the state transition matrix of the equivalent MIMO system, $\tilde{\mathbf{A}}$, is given by

$$\tilde{\mathbf{A}} = \mathbf{A}_{M-1} \mathbf{A}_{M-2} \dots \mathbf{A}_1 \mathbf{A}_0 \quad (\text{B6})$$

When the two matrices $\mathbf{A}_1$ and $\mathbf{A}_0$ are multiplied, we have

$$\begin{aligned}
\mathbf{A}_1 \mathbf{A}_0 &= \begin{bmatrix} \mathbf{A}'_1 & \mathbf{U}_1 \\ \mathbf{V} & \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{A}'_0 & \mathbf{U}_0 \\ \mathbf{V} & \mathbf{W} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}'_1 \mathbf{A}'_0 + \mathbf{U}_1 \mathbf{V} & \mathbf{A}'_1 \mathbf{U}_0 + \mathbf{U}_1 \mathbf{W} \\ \mathbf{V} \mathbf{A}'_0 + \mathbf{W} \mathbf{V} & \mathbf{V} \mathbf{U}_0 + \mathbf{W}^2 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}'_1 \mathbf{A}'_0 & \mathbf{A}'_1 \mathbf{U}_0 + \mathbf{U}_1 \mathbf{W} \\ \mathbf{V} & \mathbf{W}^2 \end{bmatrix}
\end{aligned} \tag{B7}$$

Similarly

$$\begin{aligned}
\mathbf{A}_2 \mathbf{A}_1 \mathbf{A}_0 &= \begin{bmatrix} \mathbf{A}'_2 & \mathbf{U}_2 \\ \mathbf{V} & \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{A}'_1 \mathbf{A}'_0 & \mathbf{A}'_1 \mathbf{U}_0 + \mathbf{U}_1 \mathbf{W} \\ \mathbf{V} & \mathbf{W}^2 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}'_2 \mathbf{A}'_1 \mathbf{A}'_0 + \mathbf{U}_2 \mathbf{V} & \mathbf{A}'_2 (\mathbf{A}'_1 \mathbf{U}_0 + \mathbf{U}_1 \mathbf{W}) + \mathbf{U}_2 \mathbf{W}^2 \\ \mathbf{V} \mathbf{A}'_1 \mathbf{A}'_0 + \mathbf{W} \mathbf{V} & \mathbf{V} (\mathbf{A}'_1 \mathbf{U}_0 + \mathbf{U}_1 \mathbf{W}) + \mathbf{W}^3 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}'_2 \mathbf{A}'_1 \mathbf{A}'_0 & \mathbf{A}'_2 \mathbf{A}'_1 \mathbf{U}_0 + \mathbf{A}'_2 \mathbf{U}_1 \mathbf{W} + \mathbf{U}_2 \mathbf{W}^2 \\ \mathbf{V} & \mathbf{W}^3 \end{bmatrix}
\end{aligned} \tag{B8}$$

This process continues until $\mathbf{A}_{M-1}$ is reached, then

$$\tilde{\mathbf{A}} = \mathbf{A}_{M-1} \cdots \mathbf{A}_1 \mathbf{A}_0 = \begin{bmatrix} \tilde{\mathbf{A}}' & \mathbf{\Phi} \\ \mathbf{V} & \mathbf{W}^M \end{bmatrix} \tag{B9}$$

where $\tilde{\mathbf{A}}'$, $\mathbf{\Phi}$ and $\mathbf{W}^M$ have dimensions of $N_b \times N_b$, $N_b \times (N_a - 1)$ and $(N_a - 1) \times (N_a - 1)$, respectively, and they are defined as

$$\tilde{\mathbf{A}}' = \mathbf{A}'_{M-1} \cdots \mathbf{A}'_1 \mathbf{A}'_0 \tag{B10}$$

$$\mathbf{\Phi} = \sum_{k=0}^{M-2} \mathbf{A}'_{M-1} \cdots \mathbf{A}'_{k+1} \mathbf{U}_k \mathbf{W}^k + \mathbf{U}_{M-1} \mathbf{W}^{M-1} \tag{B11}$$

$$\mathbf{W}^M = \begin{cases} \begin{bmatrix} \mathbf{0}_{(N_a-1-M) \times M} & \mathbf{I}_{(N_a-1-M)} \\ \mathbf{0}_{M \times M} & \mathbf{0}_{M \times (N_a-1-M)} \end{bmatrix}, & M < N_a - 1 \\ \mathbf{I}_{(N_a-1)}, & M \geq N_a - 1 \end{cases} \tag{B12}$$

The eigenvalues of $\tilde{\mathbf{A}}$ are the roots of the following characteristic polynomial

$$\det(\lambda \mathbf{I}_{(N_b+N_a-1)} - \tilde{\mathbf{A}}) = 0 \tag{B13}$$

where $\det$ is the determinant of a square matrix. Expanding $\det(\lambda \mathbf{I}_{(N_b+N_a-1)} - \tilde{\mathbf{A}})$ and we have

$$\det(\lambda \mathbf{I}_{(N_b+N_a-1)} - \tilde{\mathbf{A}}) = \det \begin{bmatrix} \lambda \mathbf{I}_{N_b} - \tilde{\mathbf{A}}' & -\Phi \\ \mathbf{V} & \lambda \mathbf{I}_{(N_a-1)} - \mathbf{W}^M \end{bmatrix} \quad (\text{B14})$$

Using the sparsity of $\mathbf{W}^M$, it is easy to obtain that

$$\det(\lambda \mathbf{I}_{(N_b+N_a-1)} - \tilde{\mathbf{A}}) = \lambda^{(N_a-1)} \cdot \det(\lambda \mathbf{I}_{N_b} - \tilde{\mathbf{A}}') \quad (\text{B15})$$

Denote the set of eigenvalues of $\tilde{\mathbf{A}}'$ as $\{\lambda_1, \lambda_2, \dots, \lambda_{N_b}\}$. Then the set of eigenvalues of $\tilde{\mathbf{A}}$ is

$$\left\{ \lambda_1, \lambda_2, \dots, \lambda_{N_b}, \underbrace{0, \dots, 0}_{(N_a-1) \text{ zeros}} \right\} \quad (\text{B16})$$

Therefore, the system of the M -PTV IIR filter is stable if all the eigenvalues of the matrix $\tilde{\mathbf{A}}'$ have magnitudes less than one.